Theoretical Physics III
Quantum Theory

Peter E. Blöchl

Caution! This version is not completed and may contain errors

Institute of Theoretical Physics; Clausthal University of Technology;
D-38678 Clausthal Zellerfeld; Germany;
http://www.pt.tu-clausthal.de/atp/

don't panic!

[1] To the title page: What is the meaning of ΦSX? Firstly, it sounds like "Physics". Secondly, the symbols stand for the three main pillars of theoretical physics: "X" is the symbol for the coordinate of a particle and represents Classical Physics. "Φ" is the symbol for the wave function and represents Quantum Physics "S" is the symbol for the entropy and represents Statistical Physics.

Foreword and Outlook

At the beginning of the 20th century physics was shaken by two big revolutions. One was the theory of relativity, the other was quantum theory. The theory of relativity introduces a large but finite parameter, the speed of light. The consequence is a unified description of space and time. Quantum theory introduces a small but finite parameter, the Planck constant. The consequence of quantum theory is a unified concept of particles and waves.

Quantum theory describes the behavior of small things. Small things behave radically different than what we know from big things. Quantum theory becomes unavoidable to describe nature at small dimensions. Not only that, but the implications of quantum mechanics determine also how our macroscopic world looks like. For example, in a purely classical world all matter would essentially collapse into a point, leaving a lot of light around.

Quantum mechanics is often considered difficult. It seems counterintuitive, because the picture of the world provided in classical physics, which we have taken for granted, is inaccurate. Our mind naturally goes into opposition, if observations conflict with the well proven views we make of our world and which work so successfully for our everyday experiences. Teaching, and learning, quantum mechanics is therefore a considerable challenge. It is also one of the most exciting tasks, because it enters some fundamentally new aspects into our thinking. In contrast to many quotes, it is my strong belief that one can "understand" quantum mechanics just as one can "understand" classical mechanics[2]. However, on the way, we have to let go of a number of prejudices that root deep in our mind.

Because quantum mechanics appears counterintuitive and even incomplete due to its probabilistic elements, a number of different interpretations of quantum mechanics have been developed. They correspond to different mathematical representations of the same theory. I have chosen to start with a description using fields which is due to Erwin Schrödinger, because I can borrow a number of important concepts from the classical theory of electromagnetic radiation. The remaining mystery of quantum mechanics is to introduce the particle concept into a field theory. However, other formulations will be introduced later in the course.

I have structured the lecture in the following way:

In the first part, I will start from the **experimental observations** that contradict the so-called classical description of matter.

I will then try to sketch how a theory can be **constructed** that captures those observations. I will not start from postulates, as it is often done, but demonstrate the process of constructing the theory. I find this process most exciting, because it is what a scientist has to do whenever he is faced with a new experimental facts that do not fit into an existing theory. Quantum mechanics comes in two stages, called first and second quantization. In order to understand what quantum mechanics is about, both stages are required, even though the second part is considered difficult and is often taught in a separate course. I will right from the beginning describe what both stages are about, in order to clear up the concepts and in order to put the material into a proper context. However, in order to keep the project tractable, I will demonstrate the concepts on a very simple, but specific

[2]This philosophical question requires however some deep thinking about what understanding means. Often the problems understanding quantum mechanics are rooted in misunderstandings about what it means to understand classical mechanics.

example and sacrifice mathematical rigor and generality. This is, however, not a course on second quantization, and after the introduction, I will restrict the material to the first quantization until I come to the last chapter.

Once the basic elements of the theory have been developed, I will demonstrate the consequences on some **one-dimensional examples**. Then it is time to prepare a **rigorous mathematical formulation**. I will discuss **symmetry** in some detail. In order to solve real world problems it is important to become familiar with the most common **approximation techniques**. Now we are prepared to approach physical problems such as **atoms and molecules**. Finally we will close the circle by discussing **relativistic particles** and **many particle problems**.

There are a number of textbooks available. The following citations[3] are not necessarily complete or refer to the most recent edition. Nevertheless, the information should allow to locate the most recent version of the book.

- Cohen-Tannoudji, *Quantenmechanik*[3]. In my opinion a very good and modern textbook with lots of interesting extensions.
- Schiff, *Quantum Mechanics*[1]. Classical textbook.
- Gasiorowicz, *Quantenphysik*[2].
- Merzbacher, *Quantum mechanics*[3].
- Alonso and Finn, *Quantenphysik*[1],
- Atkins, *Molecular Quantum Mechanics*[4]. This book is a introductory text on quantum mechanics with focuses on applications to molecules and descriptions of spectroscopic techniques.
- Messiah, *Quantum Mechanics*[2]. A very good older text with extended explanations. Very useful are the mathematical appendices.
- W. Nolting, *Grundkurs Theoretische Physik 5: Quantenmechanik*[4]. Compact and detailed text with a lot of problems and solutions.
- C. Kiefer *Quantentheorie* [5] Not a text book but easy reading for relaxation. [6]J.-L. Basdevant and J. Dalibard, *Quantum mechanics*. A very good course book from the Ecole Polytechnique, which links theory well upon modern themes. It is very recent (2002).

[3]Detailed citations are compiled at the end of this booklet.

Contents

“There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable.

There is another theory which states that this has already happened.”

— Douglas Adams, The Restaurant at the End of the Universe

Chapter 1

Waves? Particles? Particle waves!

Quantum mechanics is about unifying two concepts that we use to describe nature, namely that of particles and that of waves.

Fig. 1.1: Louis Victor de Broglie, 1892-1987. French physicist. Nobel price in Physics 1929 for the de Broglie relations $E = \hbar\omega$ and $p = \hbar k$, linking energy to frequency and momentum to inverse wavelength, which he postulated 1924 in his PhD thesis.[7]

- **Waves** are used to describe continuous distributions such as water waves, sound waves, electromagnetic radiation such as light or radio waves. Even highway traffic has properties of waves with regions of dense traffic and regions with low traffic. The characteristic properties of waves are that they are delocalized and continuous.

- On the other hand we use the concept of **particles** to describe the motion of balls, bullets, atoms, atomic nuclei, electrons, or in the example of highway traffic the behavior of individual cars. The characteristic properties of particles are that they are localized and discrete. They are discrete because we cannot imagine half a particle. (We may divide a bigger junk into smaller pieces, but then we would say that the larger junk consisted out of several particles.)

However, if we look a little closer, we find similarities between these concepts:

- A wave can be localized in a small region of space. For example, a single water droplet may be considered a particle, because it is localized, and it changes its size only slowly. Waves can also be discrete, in the case of bound states of a wave equation: A violin string (g: Violinensaite) can vibrate with its fixed natural vibrational frequency, the pitch, or with its first, second, etc. harmonic. The first overtone has two times the frequency of the pitch (g:Grundton), the second overtone has three times the frequency of the pitch and so on. Unless the artist changes the length of the vibrating part of the chord, the frequencies of these waves are fixed and discrete. We might consider the number of the overtone as analogous to number of particles, which changes in discrete, equi-spaced steps. This view of particles as overtones of some vibrating object is very close to particle concept in **quantum field theory**.

- If we consider many particles, such as water molecules, we prefer to look at them as a single object, such as a water wave. Even if we consider a single particle, and we lack some information about it, we can use a probability distribution to describe its whereabouts at least in an approximate manner. This is the realm of statistical mechanics.

Quantum mechanics shows that particles and waves are actually two aspects of more general objects, which I will call particle waves. In our everyday life, the two aspects of this generalized object are well separated. The common features become evident only at small length and time scales. There is a small parameter, namely the **Planck constant** h, that determines when the two aspects of particle waves are well separated and when they begin to blur[1]. In practice, the **reduced Planck constant** "hbar", $\hbar \stackrel{\text{def}}{=} \frac{h}{2\pi}$, is used[2]. Compared to our length and time scales, $\hbar$ is so tiny that our understanding of the world has sofar been based on the assumption that $\hbar = 0$. However, as one approached very small length and time scales, however, it became evident that $\hbar > 0$, and a few well established concepts, such as those of particles and waves, went over board.

A similar case, where a new finite parameter had to be introduced causing a lot of philosophical turmoil is the theory of relativity. The speed of light, c, is so much larger than the speed of the fastest object we can conceive, that we can safely regard this quantity as infinite in our everyday life. More accurate experiments, however, have found that $c < \infty$. As a result, two very different quantities, space and time, had to be unified.

Interestingly, the concept of a maximum velocity such as the speed of light, which underlies the theory of relativity, is a natural (even though not necessary) consequence of a wave theory, such as quantum mechanics. Thus we may argue that the theory of relativity is actually a consequence of quantum mechanics.

[1]"to blur" means in german "verschwimmen"

[2]The Planck constant is defined as h. This has been a somewhat unfortunate choice, because h appears nearly always in combination with $1/(2\pi)$. Therefore, the new symbol $\hbar = h/(2\pi)$, denoted reduced Planck constant has been introduced. The value of $\hbar$ is $\approx 10^{-34}$ Js to within 6%.

Chapter 2

Experiment: the double slit

The essence of quantum mechanics becomes evident from a single experiment[8, 9, 10, 11]: the double-slit experiment. The experiment is simple: We need a source of particles or waves, such as golf balls, electrons, water waves or light. We need an absorbing wall with two holes in it and behind the wall, at a distance, an array of detectors or a single detector that can be moved parallel to the wall with the holes. We will now investigate what the detector sees.

Fig. 2.1: Thomas Young, 1773-1829. English physician and physicist. Established a wave theory of light with the help of the double-slit experiment.

2.1 Macroscopic particles: playing golf

Let us first investigate the double-slit experiment for a simple case, namely macroscopic particles such as golf balls.

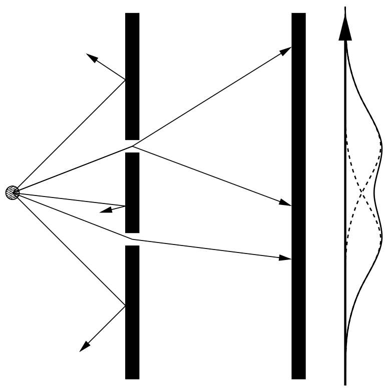


Fig. 2.2: Double-slit experiment for particles. Particles are shot randomly against a wall, which contains two slits. Some particles pass through one or the other slit in the wall and proceed to the next wall. The number of particles arriving at the second wall is counted as function of their position, yielding a probability distribution as function of their position. This probability distribution, shown as full line at the very right, is the superposition of the probability distributions, shown as dashed lines, which one obtains if one or the other slit is closed.

We take a golfer as a source of golf-balls. We build a wall with two slits. As detector we simply use little baskets that collect the golf balls that pass through the holes. By counting the number of balls in each basket that arrive per unit time, we obtain a distribution $P_{12}(y)$ as function of the basket position. We observe a distribution with two maxima, one behind each hole. If the two slits are close together the two maxima may also merge into one.

Let us repeat the experiment with the second slit closed. The golfer shots the balls at the same rate at the wall with the single remaining slit. This experiment yields a distribution of golf balls $P_1(y)$. Then we repeat the experiment with the first hole closed and obtain the distribution $P_2(y)$.

For classical particles the three distributions are not independent but they are related by

$$P_{12}(y) = P_1(y) + P_2(y) \ ,$$

that is the two distributions $P_1(y)$ and $P_2(y)$ simply add up to $P_{12}(y)$.

The explanation for this result is that the ball passes either through one or the other hole. If it passes through one hole, it does not feel the other hole, and vice versa. Hence, the probabilities simply add up.

2.2 Macroscopic waves: water waves

Now, let us turn to waves we understand well, namely water waves.

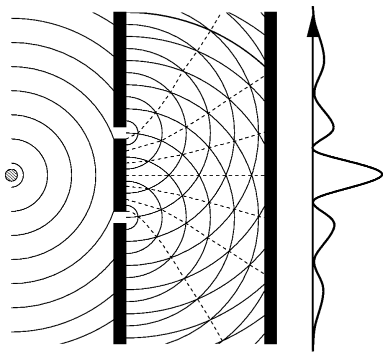


Fig. 2.3: Double-slit experiment for waves. A source on the left emits waves, which pass through two slits in a first wall. The intensity, which is proportional to the maximum squared amplitude at the second wall is monitored and shown on the very right. One observes an interference pattern for the intensity as the waves emerging from the two slits interfere. They interfere constructively in certain directions from the double slit, indicated by the dashed lines, leading to a maximum of the intensity, and they interfere destructively in other directions, leading to vanishing intensity. While the amplitudes of the waves emerging from the two slits add up to the total amplitude, the intensities obtained, when one or the other slit is closed, do not add up to the total amplitude.

The water wave is created in this experiment by a finger that periodically moves up above and down below the water surface. The frequency is characterized by the angular frequency ω.[1] The water wave moves radially outward. The height of the water surface has the form

$$\phi(\vec{r}, t) = A(|\vec{r}|) \sin\left(k|\vec{r}| - \omega t\right) = A(|\vec{r}|) \sin\left(k\left[|\vec{r}| - \frac{\omega}{k}t\right]\right) \tag{2.1}$$

It is assumed that the finger is located at the origin of the coordinate system. We see[2] that the crests and troughs of the wave move outward with constant velocity ω/k. The wave crests are separated by the wave length $\lambda = 2\pi/k$. The amplitude of the oscillations decreases with the distance from the source, because the energy of the wave is distributed over a larger region as the wave moves

[1] The angular frequency is 2π times the frequency or $\omega = 2\pi/T$ where T is the period of the oscillation.

[2] Consider a wave crest, which forms a line defined by those points for which the argument of the sinus has the value $\pi + 2\pi n$ with an integer n. Hence the line is defined by $|\vec{r}(t)| = \frac{\pi+2\pi n}{k} + \frac{\omega}{k}t$. Thus the velocity of the wave crest is $v = \frac{d|\vec{r}(t)|}{dt} = \frac{\omega}{k}$.

further away from the source[3]. This effect is described by the function $A(|\vec{r}|)$. At a given point in space, the water surface oscillates up and down with a period of $T = 2\pi/\omega$. The frequency of the oscillation is $f = \omega/(2\pi)$.

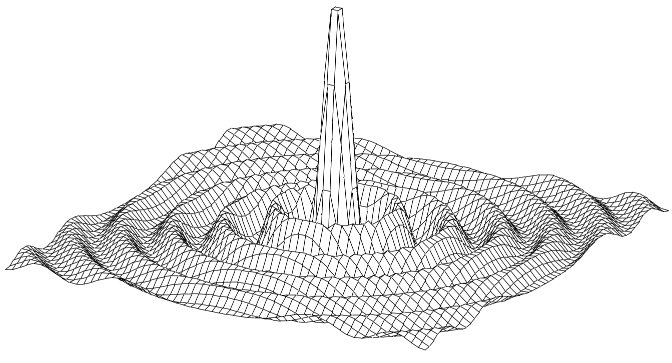


Fig. 2.4: Spherical wave

When the water wave hits the wall, it is reflected except for the two slits, where it can pass, as shown in Fig. 2.3. We have chosen the slits smaller than the wave length of the water wave. As a result, the water level inside the slits moves up and down. The slits can therefore be considered as point sources for water waves themselves. Important is that these sources are not independent, but they are synchronized. We call such waves **coherent**. Behind the wall, two waves extend in half-circles away from the two slits. Where the water waves penetrate each other, their amplitudes are superimposed. The detector observes, at each point in space, a periodic oscillation of the water surface.

Depending on the position of the detector, the waves originating from the two slits had to travel different distances to the detector. If the traveling distances from the slits to the detector differ by one-half of a wave length, one of the waves moves upward at the site of the detector, while the other wave moves downward. Since the two waves add up to zero, the detector will show no signal. If the traveling distances of the two waves are equal or if they differ by a multiple of a full wave length λ, the two waves oscillate synchronously, and, as a result, they add up to a large signal at the detector. This effect is called **interference**. As the detector is displaced parallel to the wall, it passes through regions with large and small amplitudes of oscillation of the water surface, that is through regions of **constructive interference** and **destructive interference**. Thus the distribution oscillates with the distance from the center of the two slits, as it is sketched on the right side of Fig. 2.3.

Let us make this result a little more quantitative and work out the distribution: A wave $\phi(\vec{r}, t)$

[3]The wave transports energy away from the center. Because of energy conservation, the energy flux through the surface of any surface enclosing the source must be identical. Note that, in two dimensions, this surface is a line. Since the surface area increases with distance and because the energy density must depend in some way on the amplitude, the energy-flux density must decrease with distance. The energy flux density is proportional to the square of the amplitude, and the surface, a half-circle has the length $\ell = \pi r$ is, in two dimensions, proportional to the distance r, we obtain $A(\vec{r}) \sim \frac{1}{\sqrt{\pi r}}$.

originating from a single point source, such as one of the slits, has the form

$$\phi(\vec{r},t) \overset{\text{Eq. 2.1}}{=} A(\ell)\sin(k\ell - \omega t), \tag{2.2}$$

where $\ell = |\vec{r} - \vec{r}_0|$ is the distance of the detector at position $\vec{r}$ from the point source at $\vec{r}_0$. In our experiment the two slits play the role of the point sources.

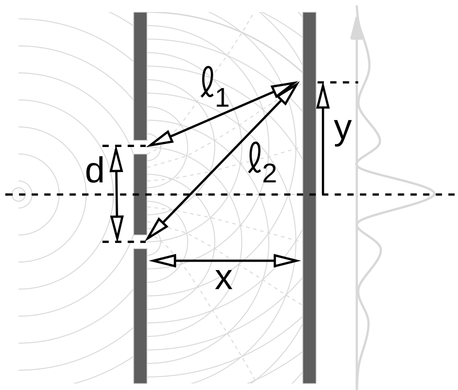


Fig. 2.5: Demonstration of the variables used for the calculation of the distribution of the double slit experiment for classical waves.

The two waves originate from slits separated by a distance d. We denote the distances of the detector from each slit as $\ell_1 = \sqrt{x^2 + (y - \frac{d}{2})^2}$ and $\ell_2 = \sqrt{x^2 + (y + \frac{d}{2})^2}$ respectively. x is the distance of the screen from the slits and y is the position on the screen relative to the projection of the mid plane of the slits. Similarly we denote the wave amplitudes as ϕ_1 and ϕ_2.

In order to design the experiment such that we can compare it with the corresponding particle experiment, we consider here the intensity of the wave instead of the amplitude.

The **intensity** is the time averaged energy flux of the wave. We have rescaled the wave amplitude such that the proportionality constant between energy flux and half of the squared wave amplitude is absorbed in the latter. Thus the intensity can be written in the form

$$I(\vec{r}) = \frac{1}{2}\Big\langle \phi^2(\vec{r},t) \Big\rangle_t = \lim_{T\to\infty} \frac{1}{T}\int_0^T dt\ \frac{1}{2}\phi^2(\vec{r},t) \tag{2.3}$$

The angular brackets denote the time average.

Let us now calculate the intensity $I_{12} = \frac{1}{2}\left\langle \frac{1}{2}(\phi_1+\phi_2)^2\right\rangle_t$ of the superposed waves from the two slits, where $\phi_1 = A(\ell_1)\sin(k\ell_1 - \omega t)$ and $\phi_2 = A(\ell_2)\sin(k\ell_2 - \omega t)$.
We introduce the mean distance $\ell_0 \stackrel{\text{def}}{=} \frac{1}{2}(\ell_1+\ell_2)$ and the difference $\Delta \stackrel{\text{def}}{=} \ell_2-\ell_1$, so that $\ell_1 = \ell_0 - \Delta/2$ and $\ell_2 = \ell_0 + \Delta/2$. Similarly we use the short hand $A_1 \stackrel{\text{def}}{=} A(\ell_1)$ and $A_2 \stackrel{\text{def}}{=} A(\ell_2)$
The sinus can be decomposed according to

$$\sin(a+b) \stackrel{\text{Eq. } O.4}{=} \sin(a)\cos(b) + \cos(a)\sin(b)\ , \tag{2.4}$$

where $a = k\ell_0 - \omega t$ and $b = \pm k\Delta/2$, to obtain

$$\sin(k\ell_0 \pm k\frac{\Delta}{2} - \omega t) \stackrel{\text{Eq. 2.4}}{=} \sin(k\ell_0 - \omega t)\cos(k\frac{\Delta}{2}) \pm \cos(k\ell_0 - \omega t)\sin(k\frac{\Delta}{2}) \tag{2.5}$$

Similarly we can write

$$\begin{aligned} A_1 &= \frac{1}{2}(A_1+A_2) + \frac{1}{2}(A_1-A_2) \\ A_2 &= \frac{1}{2}(A_1+A_2) - \frac{1}{2}(A_1-A_2) \end{aligned} \tag{2.6}$$

so that we obtain for the superposition of the two waves

$$\begin{aligned} \phi_1+\phi_2 &= A(\ell_1)\sin(k\ell_1-\omega t) + A(\ell_2)\sin(k\ell_2-\omega t) \\ &\stackrel{\text{Eq. 2.5, Eq. 2.6}}{=} (A_1+A_2)\sin(k\ell_0-\omega t)\cos(k\Delta/2) \\ &+ (A_1-A_2)\cos(k\ell_0-\omega t)\sin(k\Delta/2) \end{aligned} \tag{2.7}$$

Now we introduce the following shorthand notations R, S, α

$$R \stackrel{\text{def}}{=} (A_1+A_2)\cos(k\Delta/2) \tag{2.8}$$

$$S \stackrel{\text{def}}{=} (A_1-A_2)\sin(k\Delta/2) \tag{2.9}$$

$$\cos(\alpha) \stackrel{\text{def}}{=} R/\sqrt{R^2+S^2} \tag{2.10}$$

$$\sin(\alpha) = S/\sqrt{R^2+S^2} \tag{2.11}$$

The last relation for the sinus follows directly from the identity $\cos^2(\alpha) + \sin^2(\alpha) = 1$.

$$\begin{aligned} \phi_1+\phi_2 &\stackrel{\text{Eq. 2.7}}{=} R\sin(k\ell_0-\omega t) + S\cos(k\ell_0-\omega t) \\ &= \sqrt{R^2+S^2}\Big[\sin(k\ell_0-\omega t)\cos(\alpha) + \cos(k\ell_0-\omega t)\sin(\alpha)\Big] \\ &\stackrel{\text{Eq. 2.4}}{=} \sqrt{R^2+S^2}\sin(k\ell_0-\omega t+\alpha) \end{aligned} \tag{2.12}$$

In the last step we used again Eq. 2.4. *cont'd*

From the last equation we can read that the intensity of the oscillation averaged over time is given by the squared pre-factor $R^2 + S^2$ multiplied by a factor $1/2$ that results from the time average of the oscillatory part.

$$
\begin{aligned}
I_{12} &\overset{\text{Eq. 2.3}}{=} \lim_{T\to\infty}\frac{1}{T}\int_0^T dt\,\frac{1}{2}\left(\phi_1(t)+\phi_2(t)\right)^2 \overset{\text{Eq. 2.12}}{=} \frac{1}{4}\left(R^2+S^2\right)\\
&\overset{Eqs.\ 2.8,2.9}{=} \frac{1}{4}(A_1+A_2)^2\cos^2(k\Delta/2)+\frac{1}{4}(A_1-A_2)^2\sin^2(k\Delta/2)\\
&\overset{\text{Eq. }O.2}{=} \frac{1}{4}A_1^2+\frac{1}{4}A_2^2+\frac{1}{2}A_1A_2(\cos^2(k\Delta/2)-\sin^2(k\Delta/2))\\
&\overset{\text{Eq. }O.2}{=} \frac{1}{4}A_1^2+\frac{1}{4}A_2^2+\frac{1}{2}A_1A_2(1-2\sin^2(k\Delta/2))\\
&\overset{\text{Eq. }O.7}{=} \underbrace{\frac{1}{4}A_1^2}_{I_1}+\underbrace{\frac{1}{4}A_2^2}_{I_2}+\frac{1}{2}A_1A_2\cos(k\Delta))
\end{aligned}
$$

One factor $\frac{1}{2}$ in the proportionality between intensity and A^2 stems from the definition of the intensity and the other from the time average of the squared sinus function.

INTERFERENCE PATTERN IN THE INTENSITY

Thus we obtain the intensity $I_{12}(y) \overset{\text{Eq. 2.3}}{=} \left\langle \frac{1}{2}(\phi_1(y,t)+\phi_2(y,t))^2\right\rangle_t$ when both slits are open as

$$
I_{12}(y) = I_1(y) + I_2(y) + \underbrace{2\sqrt{I_1(y)I_2(y)}\cos(k\Delta(y))}_{\text{Interference term}} \tag{2.13}
$$

from the intensities $I_1(y) = \left\langle \frac{1}{2}\phi_1^2(y)\right\rangle_t$ and $I_2(y) = \left\langle \frac{1}{2}\phi_2^2(y)\right\rangle_t$ that are obtained, when one or the other slit is closed

$\Delta(y)$ in Eq. 2.13 is a function [4] that varies monotonously from $-d$ for $y=-\infty$ to $+d$ at $y=+\infty$ as shown in Fig. 2.6. Close to the center, that is for $y=0$ and if the screen is much farther from the slits than the slits are separated, Δ is approximately $\Delta(y) = yd/x + O(y^2)$. Hence the spatial oscillation near the center of the screen has a period [5] of $2\pi x/(kd)$.

An important observation is that the intensities do not simply add up as the probabilities do for particles. The first two terms are just the superposed intensities of the contribution of the two slits. If we interpret the intensities as probabilities for a particle to arrive, they correspond to the probability that a particle moves through the left slit and the probability that it moved to the right slit. The

[4]

$$
\begin{aligned}
\Delta(y) &= \ell_2-\ell_1 = \sqrt{x^2+(y+\frac{1}{2}d)^2}-\sqrt{x^2+(y-\frac{1}{2}d)^2} = \sqrt{x^2+y^2+\frac{d^2}{4}+dy}-\sqrt{x^2+y^2+\frac{d^2}{4}+dy}\\
&= \sqrt{x^2+y^2+\frac{d^2}{4}}\left[\sqrt{1+\frac{yd}{x^2+y^2+\frac{d^2}{4}}}-\sqrt{1-\frac{yd}{x^2+y^2+\frac{d^2}{4}}}\right]\\
&\overset{d<<x,y}{\approx} \sqrt{x^2+y^2}\left[\sqrt{1+\frac{yd}{x^2+y^2}}-\sqrt{1-\frac{yd}{x^2+y^2}}\right] \overset{\text{Taylor}}{=} \sqrt{x^2+y^2}\frac{yd}{x^2+y^2} = d\frac{y}{\sqrt{x^2+y^2}}\\
&\overset{\sin\gamma\overset{\text{def}}{=}\frac{y}{\sqrt{x^2+y^2}}}{=} \sin(\gamma)d
\end{aligned}
$$

[5] The period Λ of $\cos(k\Delta(y)$ is given by $k\Delta(y+\Lambda) = k\Delta(y)+2\pi$. In the center of the screen, i.e. $y=0$ and for small wave length, we can Taylor-expand Δ and obtain $k\frac{d\Delta}{dy}\Lambda = 2\pi$. Thus the period is $\Lambda = \frac{2\pi}{k}\left(\frac{d\Delta}{dy}\right)^{-1} = \frac{2\pi x}{kd}$.

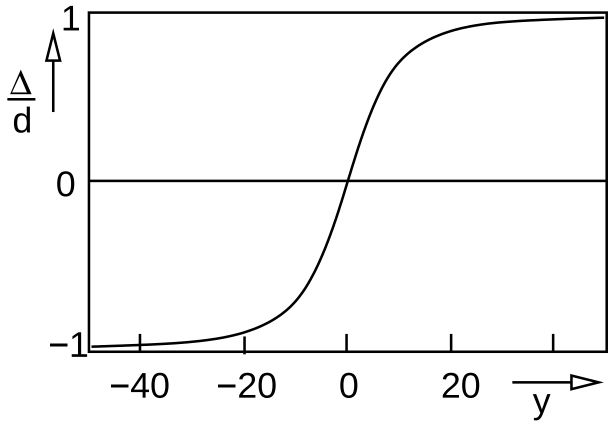


Fig. 2.6: $\Delta(y) = \ell_2 - \ell_1$ for a a double slit with $d = 1$ and a screen separated from the slit by $x = 10$. The function Δ is approximately linear $\Delta(y) \approx d/x$ in the center of the interference pattern, i.e. for $y \approx 0$, and it is nearly constant for $x >> d$. For large deflections, it saturates at a value $\Delta(|y| >> x) = \pm d$.

last, oscillatory term is a property of waves called **interference term**.

The interference term oscillates more rapidly if the wave length $\lambda = 2\pi/k$ becomes small compared to the distance of the slits, i.e. $\lambda << d$. This is shown in Fig. 2.7. In the extreme case of a very short wave length, the interference term oscillates so rapidly that the oscillations are averaged out by any experiment that has a finite resolution. In this limit, the oscillations become invisible. In such an 'inaccurate' experiment, the wave intensities behave additive just as probabilities for particles. Hence, waves with very short wave lengths behave in a certain way like particles. Waves with longer wave lengths do not.

DISAPPEARANCE OF THE INTERFERENCE TERM FOR SMALL WAVE LENGTH

In the limit of very small wave length, the interference term is averaged out in any experiment having a finite resolution. The non-oscillatory term of the intensities, which can still be observed, behaves additive just like the probability of particles. Thus the wave character of particles disappears in the limit of short wavelength.

This is a very important result, which lies at the heart of **measurement theory** and the **classical limit** of quantum mechanics that will be described later.

2.3 Microscopic waves: light

Sofar, everything sounds familiar. Now let us switch to the microscopic world and investigate light, which has a much shorter wave length:

The wave length of visible light varies from 750 nm for red light to 400 nm for blue light. If the wave length is longer than for visible light, we pass through infrared, namely heat radiation, to microwaves as they are used for radio and mobile telephony. For wave lengths shorter than that of visible light, we pass through ultraviolet light up to X-ray radiation. The wave length of visible light is still large compared to atomic dimensions. Atoms in a crystal have a typical distance of 1-2 Å or 0.1-0.2 nm, which is why X-rays need to be used to determine crystal structures.

If we perform the two-slit experiment with light, we need to use photo-detectors. A photo-detector contains electrons loosely bound to defect atoms in a crystal. When such a crystal is placed in an electric field, such as that of a light beam, the electrons can follow the applied electric field, resulting in a measurable current.

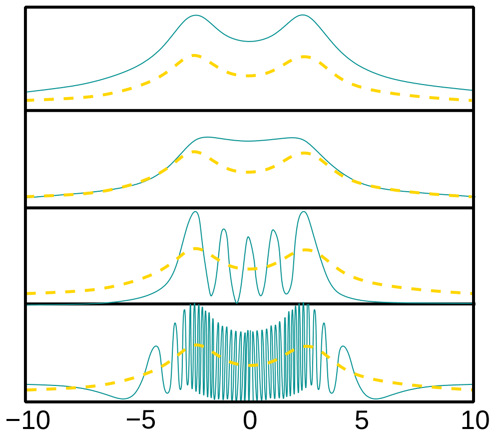


Fig. 2.7: Interference pattern of a two dimensional wave through a double slit with separation $d = 5$, on a screen located at a distance $z = 1$ behind the slits. The wave length has values $\lambda = 0.2, 2, 20, 200$ with increasing values from the bottom to the top. Each figure contains in addition the sum of intensities from the individual slits (yellow dashed line), which would correspond to a pure particle picture.

Let us now return to the double-slit experiment. For a regular light source, we obtain the same interference pattern as for water waves, only a lot smaller in its dimensions.

What happens as we slowly turn down the light? At first, the intensity measured by the photo detector becomes smaller, but the signal maintains its shape. If we turn the light further down, after the intensity is far too small to be observed with the naked eye, the signals begins to flicker and for even lower intensity random independent flashes of very low intensity are observed. If we use an array of photo detectors at different positions, only one detects a flash at a time, while the others are completely silent. (The experimentalists assured me that this is not a problem with their photo detectors.) If we add up the signals over time, the original interference pattern is observed.

Thus if the light is turned down, it behaves lumpy, like particles! These particles are called **photons**.

Apparently, if the light is strong, many photons are detected in short time intervals, so that we obtain a signal that appears continuous. In reality we see many particles. Thus our experience has misguided us, and we have introduced a concept of a continuous wave, even though we actually observe particles.

2.4 Microscopic particles: electrons

Let us now go to the microscopic world of particles. We choose electrons as example. We mount a small metallic tip so that it points towards a metallic plate that has two slits in it. Between tip and metal plate we apply a large voltage. Near the tip the electric field becomes so large, that it rips electrons out of the tip, which in turn travel towards the metallic plate. The metallic plate has two slits, which allows some electrons to pass through.

Electrons that pass through the slits are collected by a detector. We count the electrons that arrive at the detector, when the detector is at a given position. The electrons pass one by one, and enter the detector as individual particles, but when we count a large number of them as a function of position of the detector, we obtain a clear interference pattern as in the case of light! This is shown in Fig. 2.8.

We remember that the interference pattern for water waves resulted from the wave passing through both holes simultaneously. We expect that a particle, which passes through either one or the other hole cannot have an interference pattern.

Let us change the experiment. We close one hole, collect the signals, then close the other, and add up the signals. The result are two maxima, but no interference pattern. Apparently there is not only the possibility that the particle passes through one or the other slit, but also the probability that it passes through both slits simultaneously. The latter possibility makes the difference between the two humps, which are expected for particles, and the interference pattern, expected for waves.

A movie and demonstration of the double-slit experiment can be found at the Hitachi Web site `http://www.hqrd.hitachi.co.jp/em/doubleslit.cfm`

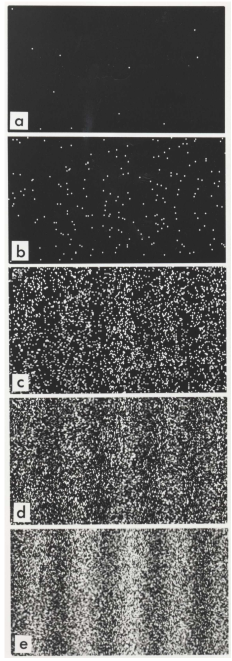


Fig. 2.8: Results of a double-slit-experiment performed by Dr. Tonomura showing the build-up of an interference pattern of single electrons. Numbers of electrons are 10 (a), 200 (b), 6000 (c), 40000 (d), 140000 (e).[12, 11, 13]

2.5 Summary

We have learned that waves such as light are composed of fixed quanta, which we may identify with particles. Particles such as electrons on the other hand are wavy in their collective behavior. In the macroscopic world, either the particle properties or the wave properties dominate, so that we have developed two different concepts for them. In reality they are two aspects of the same concept, which is the realm of quantum theory.

The observations of the double-slit experiment cannot be derived from any theory of classical waves or classical particles independently. What we observe is a new phenomenon. A new theory is required, and that theory is quantum theory. Together with Einstein's theory of relativity, the advent quantum theory changed radically our view of the world in the early 20th century.

2.6 Recommended exercises

1. Interference: Exercise 16.1 on p. 243

Chapter 3

Towards a theory

In the previous section, we learned that particle beams exhibit interference patterns as they are known from waves. To be more precise, the probability of finding a particle at a point in space exhibits an interference pattern. Therefore, we need to construct a theory that on the one hand is able to describe interference patterns on a small length scale and on the other hand it must match classical mechanics on a macroscopic length scale. One possibility is to consider particles as **wave packets**. If the size of the wave packet and the wave length is small on macroscopic length scales, we may be fooled in thinking that the particle is in fact a point. On the other hand, the wavy nature may reveal itself on a length scale comparable to the wave length.

In the following, we will try to find a theory of waves in the form of differential equations, in which wave packets behave, on a macroscopic length scale, like classical particles.

Fig. 3.1: Erwin Schrödinger, 1887-1961. Photograph from 1933. Austrian Physicist. Nobel price 1933 in physics for the Schrödinger equation.

3.1 Particles: classical mechanics revisited

Before we investigate the behavior of wave packets, let us revisit the main features of classical mechanics. One requirement of our theory is that the wave packets behave like classical particles, if the wave length is small. Therefore, we first need to have a clear understanding of the dynamics of classical particles.

In classical mechanics, the motion of particles is predicted by Newton's equations of motion

$$m\ddot{x} = F = -\partial_x V(x)$$

where each dot denotes a time derivative, and $\partial_x V$ is the derivative of the potential $V(x)$ with respect to the position x. Newton's equations relates the acceleration $\ddot{x}$ of a particle to the force F acting on the particle with the mass m as proportionality constant.

The equations of motion can be derived from another formalism, the **principle of least action**[1]. The principle of least action is particularly useful to obtain a consistent set of forces from, for example, symmetry arguments.

We start from a so-called **Lagrangian** $\mathcal{L}(v, x)$, which depends on positions x and velocities $v = \dot{x}$. For a particle in a potential, the Lagrangian has the form kinetic energy minus potential energy

$$\mathcal{L}(v, x) = \frac{1}{2}mv^2 - V(x)$$

[1]"principle of least action" translates as "Wirkungsprinzip; Hamilton principle" into German

The principle of least action says that from all conceivable paths $x(t)$ starting at time t_1 from x_1 and arriving at time t_2 at their destination x_2, the particle will choose that path which minimizes the **action** $\mathcal{S}$.[2] The action is a **functional** [3] of the path defined as follows:

ACTION

The action $\mathcal{S}$ is a functional of a path $x(t)$. It is the integral of the Lagrangian $\mathcal{L}$ over a path

$$\mathcal{S}[x(t)] = \int_{x_1,t_1}^{x_2,t_2} dt\, \mathcal{L}(\dot{x}, x) \tag{3.1}$$

The integration bounds indicate that the path begins at a specified space-time point (x_1, t_1) and ends at another one, namely (x_2, t_2). When the path is varied within the framework of the principle of least action, these boundary conditions remain unchanged.

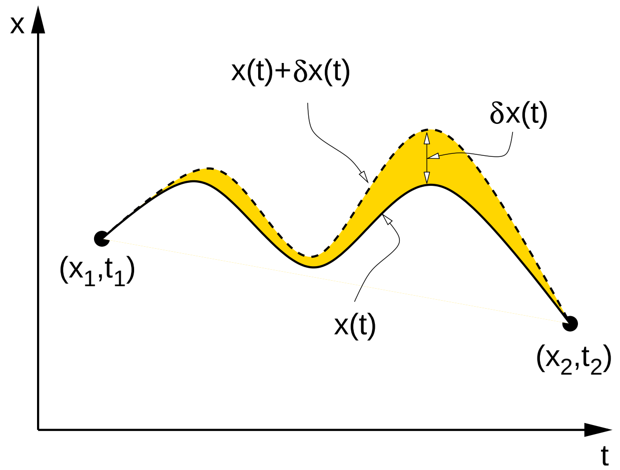


Fig. 3.2: Two paths, $x(t)$ and $x(t) + \delta x(t)$, connecting two space-time points, (x_1, t_1) and (x_2, t_2). Each path has a different value for the action $S[x(t)]$. The action depends on the path as a whole. If the action does not change to first order with respect to all possible small variations $\delta x(t)$, the action is stationary for this path, and the path is a physical trajectory from the initial to the final points.

The minimum, or better a stationary point, of the action is obtained as follows: We consider a path $x(t)$ and a small variation $\delta x(t)$. The variation $\delta x(t)$ vanishes at the initial and the final time. Now we test whether an arbitrary variation $\delta x(t)$ can change the action to first order. If that is not

[2] One should better say that the physical path makes the action stationary rather than minimize it. Historically this more general formulation has been noticed only long after the principle of least action has been formulated, so that we often speak about minimizing the action, even though this statement is too limited.

[3] A functional maps a function onto a number. To show the analogy, a function maps a number onto a number. A functional can be considered a function of a vector, where the vector index has become continuous.

the case, the action is stationary for the path $x(t)$.

$$
\begin{aligned}
\mathcal{S}[x(t)+\delta x(t)] &= \int dt\ \mathcal{L}(\dot{x}+\delta\dot{x}, x+\delta x) \\
&\stackrel{\text{Taylor}}{=} \int dt\ \Big[\mathcal{L}(\dot{x},x)+\frac{\partial\mathcal{L}}{\partial v}\delta\dot{x}+\frac{\partial\mathcal{L}}{\partial x}\delta x+O(\delta x^2)\Big] \\
&\stackrel{\text{part. int.}}{=} \mathcal{S}[x(t)]+\int dt\Big[\frac{d}{dt}\Big(\frac{\partial\mathcal{L}}{\partial v}\delta x\Big)-\Big(\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial v}\Big)\delta x+\frac{\partial\mathcal{L}}{\partial x}\delta x+O(\delta x^2)\Big] \\
&= \mathcal{S}[x(t)]+\underbrace{\Big[\frac{\partial\mathcal{L}}{\partial v}\delta x\Big]_{x_1,t_1}^{x_2,t_2}}_{=0}+\int dt\Big[-\Big(\frac{d}{dt}\frac{d\mathcal{L}}{dv}\Big)+\frac{d\mathcal{L}}{dx}\Big]\delta x(t)+O(\delta x^2) \\
\delta S &= \mathcal{S}[x(t)+\delta x(t)]-\mathcal{S}[x(t)]=\int dt\Big[-\Big(\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial v}\Big)+\frac{\partial\mathcal{L}}{\partial x}\Big]\delta x+O(\delta x^2)
\end{aligned}
$$

Thus, the functional derivative of the action is

$$
\frac{\delta\mathcal{S}}{\delta x(t)}=-\Big(\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial v}\Big)+\frac{\partial\mathcal{L}}{\partial x} \tag{3.2}
$$

The notation $O(\delta x^2)$ is a short hand notation for all terms that contains the second and higher powers of δx.

The derivation contains the typical steps, which are worth to remember:
(1) Taylor expansion of the Lagrangian to first order in coordinates and velocities.
(2) Conversion of derivatives $\delta\dot{x}$ into δx using partial integration.
(3) Removing the complete derivatives using boundary conditions $\delta x(t_1)=\delta x(t_2)=0$.

The requirement that $d\mathcal{S}=\mathcal{S}[x+\delta x]-\mathcal{S}[x]$ vanishes for any variation δx in first order of δx, translates, with the help of Eq. 3.2, therefore into the

EULER-LAGRANGE EQUATIONS

$$
\frac{d}{dt}\frac{\partial\mathcal{L}(v,x)}{\partial v}=\frac{\partial\mathcal{L}(v,x)}{\partial x} \qquad \text{with} \quad v=\dot{x} \tag{3.3}
$$

The Euler-Lagrange equations are identical to Newton's equation of motion which is easily verified for the special form $\mathcal{L}=\frac{1}{2}mv^2-V(x)$ of the Lagrangian.

3.1.1 Hamilton formalism

The **Hamilton formalism** is identical to the principle of least action, but it is formulated in more convenient terms. **Noether's Theorem**[4][14] says that every continuous symmetry results in a conserved quantity. For a system that is translationally symmetric in space and time, the two conserved quantities are momentum and energy, respectively. The Hamilton formalism is built around those two quantities.

The Hamilton formalism is obtained by a **Legendre transformation** of the Lagrangian. We define a new function, the **Hamilton function**

[4] see ΦSX: Klassische Mechanik

HAMILTON FUNCTION

$$\mathcal{H}(p,x) \stackrel{\text{def}}{=} \underset{v}{\text{stat}}\Big(pv - \mathcal{L}(v,x,t)\Big) \tag{3.4}$$

The symbol stat implies the *"stationary point"*.[a] The Hamilton function depends on positions and momenta p, but not on on the velocities! As described in the following, the velocities are considered a function $v(p,x)$ of momenta and coordinates themselves, i.e.

$$\mathcal{H}(p,x) = p \cdot v(p,x,t) - \mathcal{L}(v(p,x,t),x,t),$$

The physical meaning of the Hamilton function is the total energy.

[a] The symbol $\text{stat}_x f(x)$ implies the value of the function $f(x)$ at the point x_0, where the derivative $\partial_x f(x)$ vanishes. stat is used analogous to min or max, but it is not specific for extrema, but also saddle points having positive and negative curvatures. Caution: while the global extrema of a function are unique, a function may have several stationary points.

In order to get rid of the dependence on the velocity v in Eq. 3.4, we define the momentum p so that the Hamilton function $\mathcal{H}$ does not explicitly depend on $\dot{x}$. This leads to

$$\frac{\partial \mathcal{H}}{\partial v} = p - \frac{\partial \mathcal{L}}{\partial v} = 0$$

which defines the **canonical momentum** as follows:

CANONICAL MOMENTUM

$$p \stackrel{\text{def}}{=} \frac{\partial \mathcal{L}}{\partial v} \tag{3.5}$$

The difference between the canonical momentum described above and the **kinetic momentum** defined as $p' = mv$, that is,t as mass times velocity. The kinetic momentum is the canonical momentum of a free particle.

From Eq. 3.5 we obtain the momentum $p(x,v,t)$ as function of positions and velocities. This expression can be converted into an expression for the velocity $v(p,x,t)$ as function of momenta and positions. The latter is used in Eq. 3.4 to obtain the Hamilton function.

In the following, we derive a new set of equations of motion by forming the derivatives of the Hamilton function with respect to positions and momenta and by using the Euler Lagrange equations Eq. 3.3. The resulting equations of motion are called **Hamilton's equations**. Hamilton equations, contain exactly the same information as Newton's equation of motion and the Euler-Lagrange equations.

$$\frac{\partial \mathcal{H}}{\partial x} \stackrel{\text{Eq. 3.4}}{=} -\frac{\partial \mathcal{L}}{\partial x} + \underbrace{\left(p - \frac{\partial \mathcal{L}}{\partial v}\right)}_{=\frac{\partial \mathcal{H}}{\partial v}=0} \frac{\partial v(p,x)}{\partial x} \stackrel{\text{Eq. 3.3}}{=} -\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial v} \stackrel{\text{Eq. 3.5}}{=} -\dot{p}$$

$$\frac{\partial \mathcal{H}}{\partial p} \stackrel{\text{Eq. 3.4}}{=} v + \underbrace{\left(p - \frac{\partial \mathcal{L}}{\partial v}\right)}_{=\frac{\partial \mathcal{H}}{\partial \dot{x}}=0} \frac{\partial v(p,x)}{\partial p} = v = \dot{x}$$

We generalize the result to higher dimensions

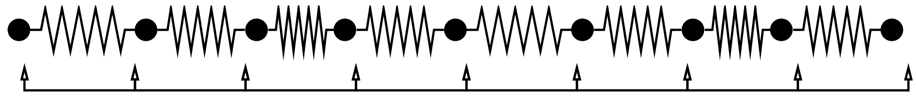
HAMILTON'S EQUATIONS OF MOTION

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i}$$
$$\dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i} \tag{3.6}$$

We will come back to the Hamilton equations in quantum mechanics, when we investigate the dynamics of wave packets. Euler-Lagrange equations, Hamilton equations and Newton's equations of motion have the same content and are mathematically identical.

3.2 Waves: the classical linear chain

Let us now remind ourselves of some properties of classical waves. As a working example, I am choosing the most simple wave I can think of, namely the vibrations of a **linear chain**. The linear chain consists of particles, or beads, distributed along a line. The beads are free to move along the direction of the line, but each bead is connected to the two neighboring beads by springs.

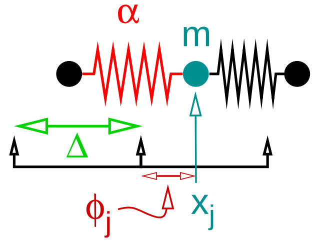


The linear chain is a simple model for lattice vibrations, where atoms oscillate about some equilibrium positions. If an atom in a crystal is displaced, the bonds to the neighboring atoms are stretched on one side and compressed on the other. The resulting forces try to restore the original atomic distances. If we let the beads loose, they will start to oscillate. These oscillations usually propagate along the entire chain. If we consider the material on a larger length scale, and if we ignore that the material is made of individual atoms, we can consider those lattice deformations as continuous waves similar to water waves or light. These lattice-distortion waves are called **phonons**. Phonons are the quanta of lattice vibrations just as photons are the quanta of light.

- Phonons are used to describe the transport of sound waves in solids.
- Phonons are responsible for heat transport in insulators. (In metals the electronic contribution to the heat transport is usually a lot larger than the lattice contribution.)

3.2.1 Equations of motion ...

Let us now describe the linear chain mathematically:



The positions of the particles are at $x_j(t) = \bar{x}_j + \phi_j(t)$. The equilibrium positions $\bar{x}_j = j\Delta$ are equally spaced with a distance Δ between neighboring particles. The displacements from the equilibrium positions are denoted by $\phi_j(t)$.

Let us write down the Lagrangian, from which we obtain the equations of motion. The Lagrangian is kinetic energy minus potential energy.

$$\mathcal{L}(\{\dot{\phi}_j\},\{\phi_j\}) = \sum_j \left[\frac{1}{2}m\dot{\phi}_j^2 - \frac{1}{2}\alpha(\phi_j - \phi_{j-1})^2\right] \tag{3.7}$$

α is the spring constant. Thus the energy of a spring is $\frac{1}{2}\alpha(d-d_0)^2$, where $d = x_j - x_{j-1}$ is the actual length of the spring and $d_0 = \bar{x}_j - \bar{x}_{j-1}$ is the length of the spring without tension.

The action is defined as the integral of the Lagrange function.

$$\mathcal{S}\left[\{\phi_j(t)\}\right] = \int_{\{\phi_{1,j}\},t_1}^{\{\phi_{2,j}\},t_2} dt\; \mathcal{L}(\{\dot{\phi}_j(t)\},\{\phi_j(t)\})$$

We use curly brackets to denote all elements in a vector, and brackets to denote functions. The action is a functional of the time-dependent displacements $[\{\phi_j(t)\}]$. According to the variational principle, the physical displacements $\phi_j(t)$ are the extremum of the action under the condition that the displacements at the initial time t_1 and the final time t_2 have fixed values.

We determine the equations of motion for the displacements $\phi_j(t)$ from the extremum condition, that is the Euler-Lagrange equations Eq. 3.3.

$$\begin{aligned}
\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\phi}_n} &\stackrel{\text{Eq. 3.3}}{=} \frac{\partial\mathcal{L}}{\partial\phi_n} \\
\stackrel{\text{Eq. 3.7}}{\Rightarrow} \frac{d}{dt}\left(m\dot{\phi}_n\right) &= \frac{\partial}{\partial\phi_n}\left(-\frac{1}{2}\alpha(\phi_n-\phi_{n-1})^2-\frac{1}{2}\alpha(\phi_{n+1}-\phi_n)^2\right) \\
\Rightarrow m\ddot{\phi}_n &= -\alpha(\phi_n-\phi_{n-1})+\alpha(\phi_{n+1}-\phi_n) \\
\Rightarrow m\ddot{\phi}_n &= \alpha\left(\phi_{n+1}-2\phi_n+\phi_{n-1}\right) \qquad (3.8)
\end{aligned}$$

The right hand side reminds of the differential quotient of the second derivative. [5]

3.2.2 ... and their solutions

The equations of motion Eq. 3.8 are those of a multidimensional harmonic oscillator, since the forces are linear in the displacements $\{\phi_n\}$. The multidimensional harmonic oscillator can be solved with the techniques described in ΦSX: Klassische Mechanik[14].

The problem Eq. 3.8 is even simpler than the multidimensional harmonic oscillator, because it is translationally invariant. Translationally invariant problems are best solved with an exponential Ansatz.

$$\phi_j(t) = \int\frac{dk}{2\pi}a(k)\mathrm{e}^{i(k\bar{x}_j-\omega(k)t)} \qquad (3.9)$$

We insert this Ansatz into the equations of motion Eq. 3.8 for the linear chain. This will provide us with the **dispersion relation** of the linear chain. A dispersion relation relates the frequency $\omega(k)$ to the wave vector k. Another form of the dispersion relation connects the energy of a particle to its momentum. Later, we will see that there is an intimate connection between energy and frequency on the one hand and the momentum and the wave vector on the other hand.

$$\begin{aligned}
m\ddot{\phi}_n &\stackrel{\text{Eq. 3.8}}{=} \alpha\left(\phi_{n+1}-2\phi_n+\phi_{n-1}\right) \\
\stackrel{\text{Eq. 3.9}}{\Rightarrow} \int\frac{dk}{2\pi}(-m\omega^2(k))a(k)\mathrm{e}^{i(k\bar{x}_n-\omega t)} &= \int\frac{dk}{2\pi}\alpha\;\underbrace{(\mathrm{e}^{ik\Delta}-2+\mathrm{e}^{-ik\Delta})}_{\left[\mathrm{e}^{ik\frac{\Delta}{2}}-\mathrm{e}^{ik\frac{\Delta}{2}}\right]^2=(2i)^2\sin^2(k\frac{\Delta}{2})}\;a(k)\mathrm{e}^{i(k\bar{x}_n-\omega t)} \\
m\omega^2(k) &= 4\alpha\sin^2(\frac{k\Delta}{2}) \\
\omega(k) &= \pm2\sqrt{\frac{\alpha}{m}}\sin(\frac{k\Delta}{2}) \qquad (3.10)
\end{aligned}$$

[5] Differential quotient of the second derivative

$$\frac{d^2f}{dx^2} = \lim_{\Delta\to0}\frac{f(x+\Delta)-2f(x)+f(x-\Delta)}{\Delta^2}$$

Eq. 3.10 is the dispersion relation of the linear chain. Its graph is shown in Fig. 3.3. Like the dispersion relation of light, the one of the linear chain behaves almost linear for small wave vectors. In contrast to light, however,the linear chain has a maximum **angular frequency**[6] of $\omega_{\text{max}} = 2\sqrt{\alpha/m}$. This is a consequence of the discrete nature of the linear chain.

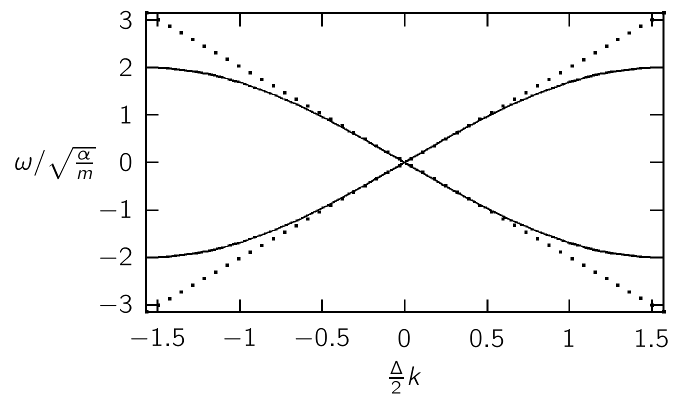


Fig. 3.3: Dispersion relation $\omega(k)$ of the discrete linear chain. The straight (dotted) lines are the dispersion relation of a continuous linear chain. The dispersion relation of the discrete chain is periodic with period $\frac{2\pi}{\Delta}$. (Each branch by itself has a period of $\frac{4\pi}{\Delta}$.)

With the dispersion relation Eq. 3.10, we can write down the solution Eq. 3.9 for the linear chain as

$$\phi_j(t) = \int \frac{dk}{2\pi} \left(a_+(k)\mathrm{e}^{i(k\bar{x}_j - |\omega(k)|t)} + a_-(k)\mathrm{e}^{i(k\bar{x}_j + |\omega(k)|t)} \right)$$

with

$$a_+(k) = a_-(-k)^*$$

The last requirement ensures that the displacements are real. [7]

Usually, we can work with each component independently of the other. The coefficients $a_\pm(k)$ are determined from the initial conditions.

$$\phi_j(t=0) = \int \frac{dk}{2\pi} \left(a_+(k) + a_-(k)\right)\mathrm{e}^{ik\bar{x}_j}$$
$$\partial_t\phi_j(t=0) = -\int \frac{dk}{2\pi}\, i|\omega(k)|\left(a_+(k) - a_-(k)\right)\mathrm{e}^{ik\bar{x}_j}$$

[6]The angular frequency ω is defined as

$$\omega = \frac{2\pi}{T} = 2\pi f \tag{3.11}$$

where T is the duration of a period and $f = 1/T$ is the frequency.

[7]Here we show that the requirement that a field has only real values implies the relation $a(k) = a^*(-k)$ of its Fourier components: Let $\phi_j = \phi(x_j)$, where $x_j = \Delta j$. The requirement that the displacements $\phi(x)$ are real has the form

$$\phi(x) = \phi^*(x) \quad \overset{\phi_j = \int \frac{dk}{2\pi}\mathrm{e}^{ikx}a(k)}{\Rightarrow} \quad \int \frac{dk}{2\pi}\mathrm{e}^{ikx}a(k) = \underbrace{\int \frac{dk}{2\pi}\mathrm{e}^{-ikx}a^*(k)}_{=\int \frac{dk}{2\pi}\mathrm{e}^{ik\Delta j}a^*(-k)}$$

$$\Rightarrow \int \frac{dk}{2\pi}\mathrm{e}^{ikx}\left(a(k) - a^*(-k)\right) = 0 \quad \Rightarrow \quad a(k) = a^*(-k)$$

q.e.d.

3.3 Continuum limit: transition to a field theory

Sofar, the linear chain has been formulated by discrete beads. For our purposes, the discrete nature of the chain is of no interest.[8] Therefore, we take the continuum limit $\Delta \to 0$ of the linear chain. In this way make the transition to a field theory.

We start out with the Lagrangian Eq. 3.7 for the discrete linear chain, and modify it introducing the equilibrium distance Δ at the proper positions without changing the result.

$$\mathcal{L}(\{\partial_t\phi_j\},\{\phi_j\}) \overset{\text{Eq. 3.7}}{=} \lim_{\Delta\to 0} \underbrace{\sum_j \Delta}_{\to\int dx}\Big[\frac{1}{2}\underbrace{\frac{m}{\Delta}}_{\to\varrho}(\partial_t\phi_j)^2(t) - \frac{1}{2}\underbrace{\alpha\Delta}_{\to\mathcal{E}}\underbrace{\left(\frac{\phi_j(t)-\phi_{j-1}(t)}{\Delta}\right)^2}_{\to(\partial_x\phi)^2}\Big]$$

As Δ approaches zero, we can convert the sum into an integral and the differential quotient will approach the derivative. We also see that a meaningful limit is obtained only if the mass is scaled linearly with Δ, that is $m(\Delta) = \varrho\Delta$ and if the spring constant is scaled inversely with Δ, that is $\alpha(\Delta) = \frac{\mathcal{E}}{\Delta}$. ϱ is the mass density of the linear chain and $\mathcal{E}$ is its **linear elastic constant**[9]. Thus the potential-energy term in the Lagrangian can be identified with the strain energy. Finally, the displacements $\phi_j(t)$ of the individual beads must be replaced by a displacement field $\phi(x,t)$ so that $\phi(x_j,t) = \phi_j(t)$ for $x_j = j\Delta$.

The limit $\Delta \to 0$ then yields

$$\begin{aligned}\mathcal{L}[\phi(x),\partial_t\phi(x)] &= \int dx\left[\frac{1}{2}\varrho\left(\partial_t\phi(x)\right)^2 - \frac{1}{2}\mathcal{E}(\partial_x\phi(x))^2\right]\\ &\overset{c\overset{\text{def}}{=}\sqrt{\frac{\mathcal{E}}{\varrho}}}{=} \frac{1}{2}\mathcal{E}\int dx\left[\frac{1}{c^2}\left(\partial_t\phi(x)\right)^2 - \left(\partial_x\phi(x)\right)^2\right] \qquad (3.12)\end{aligned}$$

where we have introduced the **speed of sound** $c = \sqrt{\frac{\mathcal{E}}{\varrho}}$. The identification of c with the speed of sound will follow later from the Euler-Lagrange equation and the corresponding dispersion relation.

The Euler-Lagrange equations for fields are only slightly more complicated compared to those for particles. First we extract the **Lagrangian density** ℓ. The spatial integral of the Lagrangian density yields the Lagrange function and the integral over space *and* time yields the action.

For the linear chain, the Lagrangian density is

$$\ell(\phi,\partial_t\phi,\partial_x\phi,x,t) = \frac{\mathcal{E}}{2}\Big[\frac{1}{c^2}(\partial_t\phi)^2 - (\partial_x\phi)^2\Big] \qquad (3.13)$$

EULER-LAGRANGE EQUATIONS FOR FIELDS

The Euler-Lagrange equations for fields have the form (see ΦSX: Klassische Mechanik[14])

$$\partial_t\frac{\partial\ell}{\partial(\partial_t\phi)} + \partial_x\frac{\partial\ell}{\partial(\partial_x\phi)} - \frac{\partial\ell}{\partial\phi} = 0 \qquad (3.14)$$

For the linear chain, we obtain the following equations of motion from Eq. 3.14

$$\frac{1}{c^2}\partial_t^2\phi(x,t) = \partial_x^2\phi(x,t)$$

[8] Our ultimate goal is to develop a field theory for particles in free space. Since space and time do not exhibit any structure, we can as well consider an infinitely fine-grained linear chain.

[9] The **elastic constant** (or Young modulus) $\mathcal{E} = \sigma/\epsilon$ is the ratio between stress σ and strain ϵ. The **strain** is the relative displacement per unit length. The **stress** is the force per area that is to be applied to a solid bar. Whereas the elastic constant of a tree-dimensional object has the unit of pressure or "energy-divided-by-volume", what we call linear elastic constant has the unit "energy-divided-by-length".

and the dispersion relation

$$\omega = \pm ck$$

If we compare this dispersion relation with that of the discrete chain in Fig. 3.3, we see that the continuum limit is identical to the $k \to 0$ limit of the discrete chain, which is the long wave length limit.

3.4 Differential operators, wave packets, and dispersion relations

Let us now investigate the dynamics of a wave packet. In order to show the full generality of the argument, we start from a general linear differential equation of a translationally invariant system in space and time.

For a translationally invariant system, the corresponding time and space coordinates do not explicitly enter the **differential operator**, but only derivatives.[10] Hence the differential equation of such a system has the form[11]:

$$Q(\partial_t, \partial_x)\phi(x,t) = 0$$

In the special case of the continuous linear chain discussed before, the function Q has the form $Q(a,b) = \frac{1}{c^2}a^2 - b^2$ so that the differential operator has the form $Q(\partial_t, \partial_x) = \frac{1}{c^2}\partial_t^2 - \partial_x^2$. [12,13]

The Ansatz $\phi(x,t) = \mathrm{e}^{i(kx-\omega t)}$ converts the differential equation of a translationally invariant system into an algebraic equation, because

$$\begin{aligned}
\partial_t \mathrm{e}^{i(kx-\omega t)} &= -i\omega \mathrm{e}^{i(kx-\omega t)} \\
\partial_x \mathrm{e}^{i(kx-\omega t)} &= +ik \mathrm{e}^{i(kx-\omega t)}
\end{aligned} \tag{3.15}$$

so that

$$0 = Q(\partial_t, \partial_x)\mathrm{e}^{i(kx-\omega t)} = Q(-i\omega, ik)\mathrm{e}^{i(kx-\omega t)}$$

Resolving the equation $Q(-i\omega, ik) = 0$ for ω yields the **dispersion relation** $\omega_\sigma(\vec{k})$, which provides the angular frequency ω as function of the wave vector k. Depending on the order[14] of the differential equation with respect to time, we may obtain one or several solutions ω for a given k. Thus the dispersion relation may have several branches, which we denote by an index σ.

In the special case of the continuous linear chain $Q(-i\omega, ik) = -\frac{\omega^2}{c^2} + k^2$, so that the dispersion relation has two branches, namely $\omega_1 = ck$ and $\omega_2 = -ck$.

The most general solution of the differential equation is then

$$\phi(x,t) = \sum_\sigma \int \frac{dk}{2\pi}\, \mathrm{e}^{i(kx-\omega_\sigma(k)t)} a_\sigma(k) \tag{3.16}$$

with some arbitrary coefficients $a_\sigma(k)$. Once the initial conditions are specified, which determine $a_\sigma(k)$, we can determine the wave function for all future times.

[10] Any dependence on x or t in the differential operator would imply that it is not invariant with respect to the corresponding translations. The differential operator is called translationally invariant, if a shifted solution of the differential equation is again a solution of the same differential operator.

[11] Any differential equation of two variables x and t can be written in the form $Q(x,t,\partial_x,\partial_t)\phi(x,t) = 0$, where $Q(x,t,u,v)$ is an arbitrary function of the four arguments. The differential operator is defined by the power series expansion, where u is replaced by ∂_x and v is replaced by ∂_t

[12] Here we use a strange concept, namely a function of derivatives. The rule is to write out the expression as if the derivative were a number, but without interchanging any two symbols in a product, because $\partial_x x \neq x\partial_x$. Note that $\partial_x x = 1 + x\partial_x$, because in operator equations we should always consider that this operator is applied to a function, i.e. $\partial_x x f(x) = f(x) + x\partial_x f(x) = (1 + x\partial_x)f(x)$. Any differentiation occurring in a product acts on all symbol multiplied to it on its right-hand side.

[13] The operator $\Box \stackrel{\text{def}}{=} \frac{1}{c^2}\partial_t^2 - \vec{\nabla}^2$ is called the **d'Alambert operator**. The **Laplacian** is defined as $\Delta \stackrel{\text{def}}{=} \vec{\nabla}^2$.

[14] The order of a differential equation is the highest appearing power of a derivative in a differential equation.

Wave packets

Let us now investigate the behavior of a **wave packet**. We select an envelope function $\chi(x, t = 0)$, a wave vector k_0 and one particular branch, i.e. σ, of the dispersion relation.

WAVE PACKET

A wave packet has the form

$$\phi(x,t) = \mathrm{e}^{i(k_0x-\omega_\sigma(k_0)t)}\chi(x,t) \tag{3.17}$$

An example is shown in Fig. 3.4. We call $\chi(x, t)$ the **envelope function**. The envelope function modulates a plane wave. k_0 is the dominating wave vector. We usually assume that the envelope function is smooth and extended over several wave lengths $\lambda = 2\pi/|k_0|$ of the plane wave part.

We want to investigate the behavior of the dynamical evolution of an envelope function $\chi(x, t)$. By equating

$$\mathrm{e}^{i(k_0x-\omega_\sigma(k_0)t)}\chi(x,t) \stackrel{\text{Eq. 3.17}}{=} \phi(x,t) \stackrel{\text{Eq. 3.16}}{=} \int \frac{dk}{2\pi}\mathrm{e}^{i(kx-\omega_\sigma(k)t)}a_\sigma(k)$$

we obtain a expression of $\chi(x, t)$

$$\chi(x,t) = \int \frac{dk}{2\pi}\mathrm{e}^{i\left([k-k_0]x-[\omega_\sigma(k)-\omega_\sigma(k_0)]t\right)}a_\sigma(k) \tag{3.18}$$

This is an exciting result: For a system, that is translationally invariant in space and time, the dispersion relation is sufficient to predict the complete future of a wave packet.

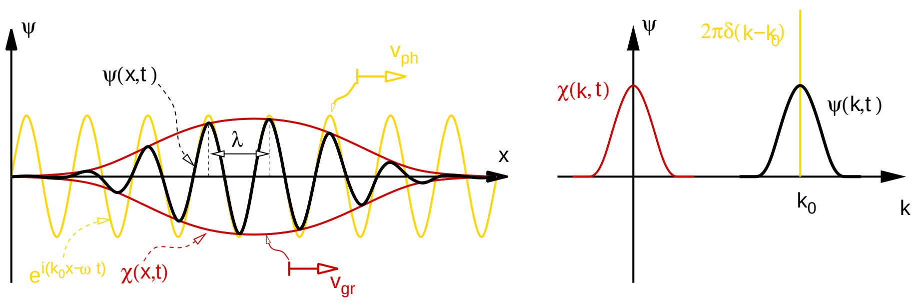


Fig. 3.4: Wave packet in real and reciprocal space. In real space, a wave packet $\Psi(x, t)$ is the product between a plane wave $\mathrm{e}^{i(kx-\omega t)}$ and a smooth envelope function $\chi(x, t)$. The plane wave moves with the phase velocity v_{ph}, while the envelope function moves with the group velocity v_{gr}. In reciprocal space, the plane wave part is a δ-function times 2π centered at the dominant wave vektor k_0. Because the envelope function is smooth, it contributes only near $k = 0$. The wave packet $\Psi(k, t)$ is a convolution $\Psi(k,t) = \int \frac{dk'}{2\pi} 2\pi\delta(k-k_0-k')\chi(k',t) = \chi(k-k_0,t)$.

Sofar, everything has been exact. Now, we need to introduce an approximation that is only valid for very smooth envelope functions $\chi(x, t)$. We assume that the envelope function is smooth on the length scale of the wave length $\lambda = 2\pi/k$. This implies that its Fourier transform $a_\sigma(k)$ has large components only for k close to $|k_0|$.[15] In this region, we can replace $\omega(k)$ by its Taylor expansion[16]

[15] If $f(x) = \int \frac{dk}{2\pi} f(k)\mathrm{e}^{ikx}$ then we find by variable substitution $x \to \alpha x$ and $k \to k/\alpha$ that $f(\alpha x) = \int \frac{dk}{2\pi}\frac{1}{\alpha}f(\frac{k}{\alpha}x)\mathrm{e}^{ikx}$. If α becomes smaller $f(\alpha x)$ becomes smoother in real space and more localized near $k = 0$ in k-space.

[16] The notation $O(x^n)$ is a shorthand notation for all terms containing powers of x of order n and higher.

$\omega(k) \approx \omega(k_0) + \frac{d\omega}{dk}|_{k_0}(k-k_0) + O(k-k_0)^2$ to first order in $k - k_0$.

$$\chi(x,t) \stackrel{\text{Eq. 3.18}}{\approx} \int \frac{dk}{2\pi} \mathrm{e}^{i[k-k_0](x-\frac{\partial\omega_\sigma}{\partial k}t)} a_\sigma(k) \stackrel{\text{Eq. 3.18}}{=} \chi(x - \underbrace{\frac{\partial\omega}{\partial k}}_{v_g} t, 0) \tag{3.19}$$

Thus, the wave packet moves with the so-called group velocity $v_g = \frac{\partial\omega}{\partial k}$, without much of a change of its shape.

Let us generalize the result to three dimensions.

GROUP VELOCITY

The wave packet as a whole, i.e. its envelope function, moves with velocity

$$\vec{v}_g \stackrel{\text{Eq. 3.19}}{=} \vec{\nabla}_{\vec{k}}\Big|_{\vec{k}_0} \omega(\vec{k}) \qquad \text{or} \qquad v_{g,i} = \frac{\partial\omega}{\partial k_i}\Big|_{\vec{k}_0} \tag{3.20}$$

It is not quite true that the wave packet does not change its shape. If we violate the assumption that the wave packet is very extended and consider envelope functions that are less smooth, then the Fourier transform is no more localized around k_0, and higher order terms of the Taylor expansion of $\omega_\sigma(\vec{k})$ become important. These terms contribute to a delocalization, or spreading, of the wave packet. This effect can be thought of in the following way: The wave packet consists of components with different velocities, so that faster components run ahead and the slower components lag behind the rest of the wave packet. As a result, the wave packet usually broadens with time. This broadening of the wave packet is called **dispersion**.

In general, the group velocity is distinct from the **phase velocity**. The phase velocity is derived from a single plane wave

$$\mathrm{e}^{i(kx-\omega t)} = \mathrm{e}^{ik(x-\overbrace{\frac{\omega}{k}}^{v_p}t)} \tag{3.21}$$

The phase velocity is the velocity of a single wave crest or wave trough.

Again, we list here the three-dimensional expression for the phase velocity

PHASE VELOCITY

$$\vec{v}_p \stackrel{\text{Eq. 3.21}}{=} \frac{\omega(\vec{k}_0)}{|\vec{k}_0|}\vec{e}_{k_0} = \frac{\omega(\vec{k}_0)}{|\vec{k}_0|}\frac{\vec{k}_0}{|\vec{k}_0|} \tag{3.22}$$

where $\vec{e}_k = \vec{k}/|\vec{k}|$ is the unit vector pointing along the wave vector $\vec{k}$. The phase velocity is the velocity of a wave crest. It can be read from the dispersion relation $\omega(k)$ as the slope of the line connecting the origin with a particular point at the dispersion relation.

The dispersion relation of the continuous linear chain is purely linear, which implies that the waves have a constant uniform velocity. This velocity is the velocity of sound, and it is analogous to the speed of light. A medium with a linear dispersion relation is convenient for signal transmission because, given a constant value for ω/k, a wave packet does maintain its shape during time even if it is localized in space. Hence, the information encoded in the shape of the wave packet remains intact during transmission.

3.5 Breaking translational symmetry

The wave packets discussed so-far move with a constant speed. This is a consequence of the translational symmetry we have imposed. In order to explore, how we can teach wave packets to

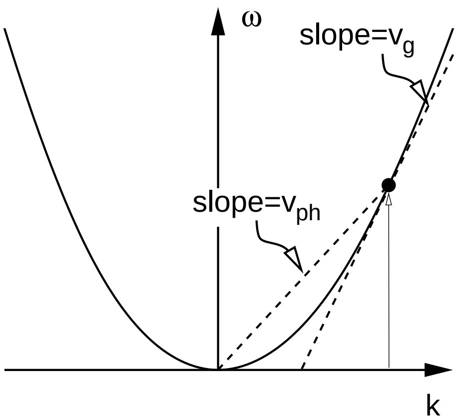


Fig. 3.5: Group and phase velocity in a dispersion relation $\omega(\vec{k})$ for a classical particle. The group velocity is the slope of the tangent of the dispersion relation, while the phase velocity is the slope of the line connecting the origin of the graph with a given point on the dispersion relation.

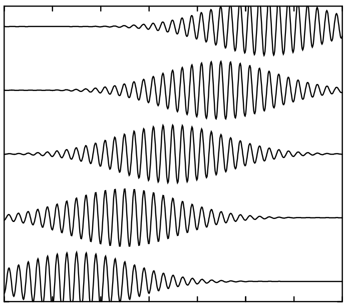


Fig. 3.6: The wave function of a linear chain. The wave packets obtained for subsequent times have been displaced upwards with increasing time. The vertical displacement for each wave packet corresponds to the longitudinal displacement of the chain-beads. The tic-marks on the x-axis denote units of the chain. The functions have been obtained by numerical solution of the differential equation for each chain bead.

behave like particles, we need to understand their behavior when they respond to external forces.

What happens, if the translational symmetry is slightly broken, such as in the presence of an external potential? In that case, the Lagrangian and the equation of motion, i.e. the differential operator $Q(\partial_t, \partial_x, x, t)$, contain terms that depend explicitly on the space coordinate.

To simplify the discussion, we assume that these terms are, on a length scale of the wave length, only slowly varying. The dispersion relation $\omega(k)$ can be written down rigorously only in a translationally symmetric situation. Without translational symmetry in space and time there is no one-to-one correspondence between frequency and wave vector. However, if we allow only slow spatial variations, we can still evaluate a spatially dependent $\omega(k,x)$ by fixing the spatially dependent terms at their local values: at each point in space, we compare to a translationally invariant system with the same local dispersion relation.

Let us define the position of the wave packet as[17].

$$x_0(t) \stackrel{\text{def}}{=} \frac{\int dx\, |\phi(x,t)|^2 x}{\int dx\, |\phi(x,t)|^2} \tag{3.23}$$

The mean k-vector is denoted by k_0. The velocity of the wave packet is the group velocity

$$\partial_t x_0(t) \stackrel{\text{Eq. 3.20}}{=} \frac{\partial}{\partial k}\,\omega\left(x_0(t), k_0(t)\right). \tag{3.24}$$

Since we still have translational symmetry in the time variables, the value of the conjugate variable $\omega(x_0(t), k_0(t))$ remains constant[18], so that

$$0 = \frac{d}{dt}\omega(x_0(t), k_0(t)) = \frac{\partial\omega}{\partial x}\underbrace{\partial_t x_0}_{\frac{\partial\omega}{\partial k}} + \frac{\partial\omega}{\partial k}\partial_t k_0 \stackrel{\text{Eq. 3.20}}{=} \frac{\partial\omega}{\partial k}\Big(\frac{\partial\omega}{\partial x} + \partial_t k_0\Big)$$
$$\Rightarrow \partial_t k_0 = -\frac{\partial}{\partial x}\omega(x_0(t), k_0(t)) \tag{3.25}$$

Let us take a closer look at what we have got: the wave packet travels according to

$$\partial_t x_0 \stackrel{\text{Eq. 3.24}}{=} \frac{d\omega}{dk} \quad ; \quad \partial_t k_0 \stackrel{\text{Eq. 3.25}}{=} -\frac{d\omega}{dx} \tag{3.26}$$

This sounds familiar! Let us compare this result with **Hamilton's equation of motion**, Eq. 3.6, for classical particles

$$\partial_t x = \frac{d\mathcal{H}}{dp} \quad ; \quad \partial_t p = -\frac{d\mathcal{H}}{dx} \tag{3.27}$$

The two sets of equations, Eq. 3.26 and Eq. 3.27, are identical, if we perform the replacement

$$\mathcal{H}(p,x) \to \hbar\omega(x, \hbar k)$$
$$p \to \hbar k$$

where $\hbar$ is some arbitrary constant, which we will later identify with the **reduced Planck constant**. (We have dropped here the subscripts of x and k.)

What we have done sofar, is to derive the equations of motion for wave packets from a specified differential operator. We have seen that the dynamics of wave packets is identical to that of classical particles, if the dispersion relations are related.

We are almost there!

[17] Consider $P(x) = \frac{|\phi(x)|^2}{\int dx |\phi(x)^2|}$ as a density. Then the definition of the local position is the average position.

[18] I have not really explained why ω is conserved with time. Let us assume a differential equation $Q(\partial_t, \partial_x, x)\Psi(x,t) = 0$. We can then make an Ansatz $\Psi(x,t) = \mathrm{e}^{-i\omega t}\phi(x,\omega)$. Inserting the Ansatz into the differential equation yields $Q(-i\omega, \partial_x, x) = 0$. For each value of ω, we obtain one or more solutions $\phi_\sigma(x,\omega)$. The general solution can therefore be written as $\Psi(x,t) = \sum_\sigma \int \frac{d\omega}{2\pi}\phi_\sigma(x,\omega)a_\sigma(\omega)\mathrm{e}^{i\omega t}$ with some arbitrary coefficients a_σ. Thus if we choose the coefficients so that only a small ω-region contributes, this region will be the same for all times. Hence, ω does not change in time.

CORRESPONDENCE PRINCIPLE

In order to construct wave packets that behave like classical particles, we start from a Lagrangian $\mathcal{L}$ for the classical particle. From the Lagrangian we derive the Hamilton function $\mathcal{H}$. The equation $E - \mathcal{H}(p, x) = 0$ is converted with the help of Eq. 3.15 into a differential equation by replacing

$$E \rightarrow \hbar\omega \rightarrow i\hbar\partial_t$$
$$p \rightarrow \hbar k \rightarrow \frac{\hbar}{i}\partial_x \tag{3.28}$$

$\hat{H} \stackrel{\text{def}}{=} \mathcal{H}(\frac{\hbar}{i}\partial_x, x)$ is called the Hamilton operator. The correspondence principle was invented by de Broglie in his PhD thesis. It is also called de Broglie relation[7].
The correspondence principle is used to translate the classical equations of motion into the quantum mechanical wave equations and vice versa.

The correspondence principle, when applied to the classical equation $E = \mathcal{H}(p, x)$, directly leads to the

SCHRÖDINGER EQUATION [15, 16, 17, 18]

$$Q(\partial_t, \partial_x, x, t)\psi(x, t) = [i\hbar\partial_t - \mathcal{H}(\frac{\hbar}{i}\partial_x, x, t)]\psi(x, t) = 0 \tag{3.29}$$

I introduced a new symbol, namely $\psi(x, t)$ instead of $\phi(x, t)$, in order to distinguish the displacement field ϕ of a linear chain from the Schrödinger wave function ψ .[19]

We found a theory that is capable of describing the interference patterns, while reproducing classical mechanics for wave packets on a macroscopic length scale. What we have done sofar is to quantize the motion of a classical particle with a given Lagrangian. This is called **first quantization**. The Schrödinger equation is for quantum mechanics what Newton's law is for classical mechanics. Most of this course will be devoted to the Schrödinger equation, how to understand it, and how to solve it.

Without proof, I add here the form of the corresponding Lagrangian, that produces the Schrödinger equation as the corresponding Euler-Lagrange equation.

$$\mathcal{L}[\psi, \partial_t\psi, \partial_x\psi, x, t] = \int dx\ \psi^*(x, t)\Big[i\hbar\partial_t - \mathcal{H}(\frac{\hbar}{i}\partial_x, x, t)\Big]\psi(x, t) \tag{3.30}$$

It is a simple exercise to show that the action for a Hamilton function $\mathcal{H} = p^2/(2m) + V(x)$ is real, and, to show, via Eq. 3.14, that the Euler-Lagrange equation for Eq. 3.30 is indeed the Schrödinger equation.

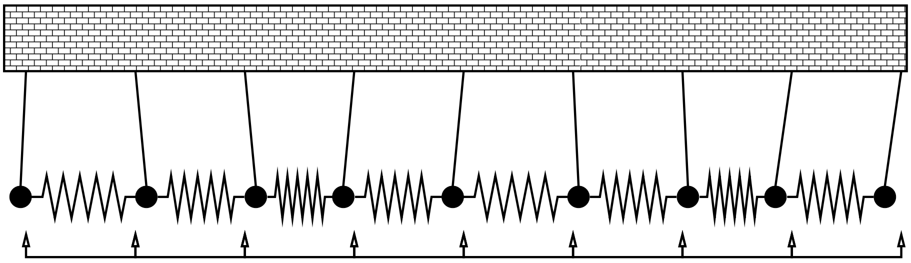
In most practical cases, the translation from a classical system in to the quantum description is unique. However, there are exceptions. In classical physics we need not pay attention to the order in which momenta and positions appear. When we replace the momentum and energy by its differential operators, however, the product px results in a different differential equation than the product xp because $\partial_x x \neq x\partial_x$. There is no unique recipe that shows us how to arrive at the correct wave-description from a classical theory alone. Experiments provided guidance to set up the Hamilton function for a classical system. When the resulting Hamiltonian is not sufficient to determine the quantum mechanical Hamilton operator, we need to resort to experiments again.

[19]The field ϕ of the hanging linear chain will be analogous to that of a relativistic particle. In the non-relativistic limit, the wave function ϕ does not directly turn into the Schrödinger field $\psi(x, t)$. An additional transformation is required to get rid of the energy related to the rest-mass.

3.6 Introducing mass into the linear chain (Home study)

We have ventured far from the linear chain, with which we began this chapter. The wave packets of the linear chain do not seem to have too much to do with particles. The reason is that the dispersion relation of the linear chain is similar to that of **photons**, the mass-less light-particles. However, it does not seem to be able to describe particles with mass such as electrons.

Here, I will show that a small variation of the linear chain actually leads to to the relativistic dispersion relation of particles that have a mass. While being a bit of a curiosity, the comparison of a purely classical system that is analogous to a quantum mechanical wave description may be very useful to imagine quantum mechanical principles. Furthermore, it may even guide the interested reader into the fundamental principles underlying the Theory of Relativity.



Let us again have a look at the continuous linear chain. In Eq. 3.12 (p. 34) we obtained the Lagrangian, which had the form

$$\mathcal{L}[\phi(x),\dot{\phi}(x)]=\frac{\mathcal{E}}{2}\int dx\Big[\frac{1}{c^2}|\partial_t\phi|^2-|\partial_x\phi|^2\Big]$$

where $\mathcal{E}$ is the linear elastic constant. There is a simple generalization of that Lagrangian. The properties we want to preserve are

- spatial and temporal translational symmetry and the so-called
- gauge symmetry[20] [21]. Gauge symmetry means that the physics depends only on the absolute value of ϕ and not its phase. Furthermore, we require that the Hamiltonian has
- only terms up to quadratic order in ϕ in order to maintain its simple form.

There is exactly one additional term, which we can include, namely a constant multiplied with the absolute square of ϕ. We choose this constant equal to $\frac{m_0^2c^2}{\hbar^2}$, where m_0 is a free parameter. We could have chosen any other prefactor. The explicit form chosen will be motivated later, when we compare the dispersion relation for the hanging linear chain with that of a relativistic particle.

$$\mathcal{L}[\phi(x),\dot{\phi}(x),\partial_x\phi(x)]=\frac{\mathcal{E}}{2}\int dx\Big[\frac{1}{c^2}|\partial_t\phi|^2-|\partial_x\phi|^2-\Big(\frac{m_0c}{\hbar}\Big)^2|\phi|^2\Big] \qquad (3.31)$$

The Euler-Lagrange equation Eq. 3.14 for this Lagrange function is the **Klein-Gordon equation**

KLEIN-GORDON EQUATION

$$\Big[\frac{1}{c^2}\partial_t^2-\partial_x^2+(\frac{m_0c}{\hbar})^2\Big]\phi(x,t)=0 \qquad (3.32)$$

[20] "gauge symmetry" translates as "Eichsymmetrie" into German
[21] Gauge is pronounced "geidsch".

If we apply the correspondence principle Eq. 3.28 backwards to the Klein-Gordon equation, we obtain the dispersion relation of a relativistic particle with rest mass m_0.

DISPERSION RELATION OF RELATIVISTIC PARTICLES

$$E^2 = p^2c^2 + m_0^2c^4,$$

Now we can identify the constants in the Klein-Gordon equation.

- c is the speed of light. This is the maximum speed $c = v_{\max} = \max \frac{\partial E}{\partial p} = c$
- m_0 is the rest mass of the particle. For non-relativistic velocities the dispersion relation is $E = m_0c^2 + \frac{p^2}{2m_0}$. Thus, in this limit, m_0 is the known inertia mass. m_0c^2 is the energy related to the mere existence of the particle. If a particle collides with its antiparticle, this energy will be turned into radiation energy.

What is the relation between a wave-packet of the Klein-Gordon equation and a relativistic particle? A wave packet of the Klein-Gordon equation cannot move faster than the speed of light c[22], which is the basis for the theory of relativity. Historically, the theory of relativity indeed has its roots in a wave theory, namely electrodynamics: Lorentz observed that the Maxwell's equations had certain symmetries, that were incompatible with the concept of an infinite speed of light. It took Einstein to realize the radical implications of this symmetry and to propose that the same symmetries are valid for the dynamics of particles.

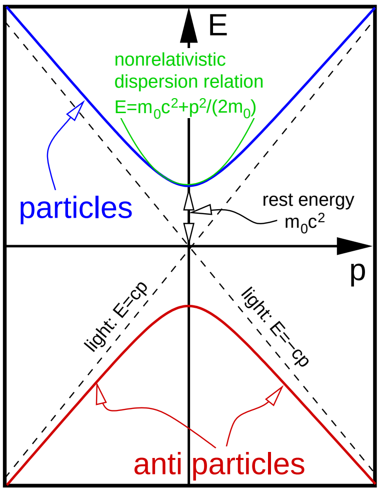


Fig. 3.7: Relativistic dispersion relation. The relativistic dispersion relation approaches $E = cp$ for large momenta, indicating that a particle cannot become faster than the speed of light. The nonrelativistic dispersion relation is obtained as Taylor expansion of $E(p)$ at $p = 0$. The relativistic dispersion relation predicts the existence of anti particles. Even at rest, a particle has an energy the rest energy. If for example, an electron and a positron, its anti particle, collide, they annihilate each other. The rest energy is emitted in the form of two photons leaving in opposite directions.

We obtained the linear chain as a purely classical model of a chain of beads connected by springs. The mechanical analogue of the Klein-Gordon equation is the hanging linear chain: A chain of pendulums connected by springs. The first term in the Klein-Gordon equation Eq. 3.32 is the inertia

[22]The velocity is $v = \partial E/\partial p = pc^2/\sqrt{p^2c^2 + m_0^2c^4} = c[1 + (\frac{m_0c}{p})^2]^{-\frac{1}{2}} \rightarrow c$ for $p \rightarrow \infty$

force of the individual beads. The second term is the force due to the springs between neighboring beads. The last term can be described by a force between the displaced bead and its equilibrium position.

Thus we have found a purely classical model, that describes a quantized relativistic particle. There is only one caveat: the displacements of the linear chain must be real, whereas the wave functions belonging to one branch of the dispersion relation are always complex.

3.7 Measurements

The purpose of any theory is to make predictions for measurable quantities. A measurable quantity is called an **Observable**. Examples for observables are energy, momentum, position etc. However there are also quantities that are not observable: The vector potential $\vec{A}(\vec{r}, t)$ of electrodynamics is such an example: A gauge transformation yields different vector potentials, which describe exactly the same physics. Thus the vector potential cannot be measured. Nevertheless, the magnetic field, which is directly related to the vector potential, can be measured.

In classical mechanics, the state of a system is described by the positions and momenta of all particles. Thus all observables can be expressed as functions $A(p, x)$ of positions and momenta. An example is the energy, which is expressed by the Hamilton function $H(p, x)$, or the angular momentum defined as $\vec{L}(\vec{r}, \vec{p}) = \vec{r} \times \vec{p}$.

Quantum mechanical wave packets, however, are spread over some volume. Therefore, they do not have a precise position nor a precise momentum. Thus, we need to perform an average in order to attribute a value to an observable for a given wave packet.

We relate the square of the wave function to an intensity, just as in optics. On the other hand, the intensity has been related to the density of particles or the probability density to find a particle at a given point in time and space.

PHYSICAL MEANING OF THE WAVE FUNCTION

The absolute square $\phi^*(x)\phi(x)$ of the wave function is interpreted as probability density for the particle to be at a given position.

As a consequence of this probabilistic interpretation of the wave function, we need to require that a intensity of a wave function integrates to unity. We call this condition the **normalization condition**.

NORMALIZATION CONDITION

$$\int dx\ \phi^*(x, t)\phi(x, t) = 1 \tag{3.33}$$

The normalization condition reflects that the probability for the system being in any state is equal to one.

Most of the previous discussion has been based on wave packets of the form $\phi(x, t) = \chi(x, t)e^{i(kx-\omega t)}$, which are very localized on a macroscopic length scale, but delocalized on a length scale set by the wave length. Intuitively we identify the position $\langle x\rangle$ of a wave packet with the center of the intensity.

$$\langle x\rangle = \int dx\ \phi^*(x)x\phi(x) = \int dx\ |\chi(x)|^2 x \tag{3.34}$$

Thus we have defined what we mean with the position of a wave packet.

Now let us turn to the momentum. The momentum of an extended wave packet should be related to the wave vector k_0 of the plane wave part.

The correspondence principle identified $\hbar k$ with the momentum of a classical particle. We can obtain the momentum of a wave packet as

$$
\begin{aligned}
\langle p\rangle &\stackrel{\text{Eq. 3.28}}{=} \int dx\, \phi^*(x)\frac{\hbar}{i}\partial_x\phi(x) \stackrel{\text{Eq. 3.17}}{=} \hbar k - i\hbar\int dx\, \chi^*(x)\partial_x\chi(x)\\
&= \hbar k - \frac{i\hbar}{2}\Big(\int dx\, \chi^*(x)\partial_x\chi(x) + \underbrace{\int dx\, \chi(x)\partial_x\chi^*(x)}_{(A)}\Big)\\
&\quad -\frac{i\hbar}{2}\Big(\int dx\, \chi^*(x)\partial_x\chi(x) - \underbrace{\int dx\, \chi(x)\partial_x\chi^*(x)}_{(A)}\Big)\\
&= \hbar k - \frac{i\hbar}{2}\int dx\, \partial_x\Big(\chi^*(x)\chi(x)\Big) - m\int dx\, \underbrace{\frac{i\hbar}{2m}\Big(\chi^*(x)\partial_x\chi(x) - \chi(x)\partial_x\chi^*(x)\Big)}_{J_\chi(x)}\\
&= \hbar k - \frac{i\hbar}{2}\underbrace{\Big[\chi^*(x)\chi(x)\Big]_{-\infty}^{\infty}}_{=0} + \hbar\int dx\, \mathrm{Im}[\chi^*(x)\partial_x\chi(x)]\\
&\stackrel{\mathrm{Im}(\chi)=0}{=} \hbar k
\end{aligned}
$$

The assumptions are that the envelope function is purely real, i.e. $\mathrm{Im}(\chi) = 0$, and that it vanishes at infinity.[23]

Thus we have defined what we mean with the momentum of a wave packet. We have used here the Symbol $\langle A\rangle$ to denote the quantum mechanical mean value of an observable A consistent with the expression Eq. 3.34 for the positions. Furthermore we required that the wave packet be localized, so that its amplitude vanishes at infinity.

Now, we need to generalize these results to more general observables A. In classical mechanics, an observable can always be expressed by a function $A(x, p)$ of the positions and momenta. This is because positions and momenta characterize a classical state uniquely, two generalizations are possible:

- we could attribute the observable to the function
$$\langle A\rangle = A(\langle p\rangle, \langle x\rangle) \tag{3.35}$$
of the mean values for positions and momenta obtained above.
- we could attribute the observable to the differential operator $A(\frac{\hbar}{i}\partial_x, x)$ and evaluate the expectation value as
$$\langle A\rangle = \int dx\, \phi^*(x)A(\frac{\hbar}{i}\partial_x, x)\phi(x) \tag{3.36}$$
similar to the expressions for the positions and momenta.

Only the second choice, Eq. 3.36, is correct. Using the former choice, Eq. 3.35, a variable transform would result in different expressions for the expectation value.

Another indication in favor of the second choice Eq. 3.36 is the following: The correspondence principle, Eq. 3.28 on p. 40, provides us, together with the Schrödinger equation Eq. 3.29, with a prescription for the energy in the form:

$$\langle E\rangle \stackrel{\text{Eqs. 3.28,3.36}}{=} \int dx\, \phi^*(x)i\hbar\partial_t\phi(x) \stackrel{\text{Eq. 3.29}}{=} \int dx\, \phi^*(x)H(\frac{\hbar}{i}\partial_x, x)\phi(x)$$

[23]The expression J_χ resembles the expression for the flux density. that is derived later in Section 8.4 on p. 133. As it is applied only to the envelope function it has a different meaning in this case, and with the assumption of a real envelope function the expression for J_χ vanishes altogether.

Since the energy $H(p, q)$ is a perfectly allowed measurable quantity, our generalization must be consistent with the expression given above. The alternative rule using $\langle A \rangle = A(\langle p \rangle, \langle x \rangle)$ would give an inconsistent result.

Therefore, we set up the general rule for the evaluation of observables: start from a classical observable $A(p, x, t)$ and replace momenta by its differential operator according to the correspondence principle.

EXPECTATION VALUE OF A GENERAL OBSERVABLE

The expectation value of an observable, which is represented by a function of momenta and position $A(p, x)$, is obtained for a quantum mechanical system from its wave function as

$$\langle A \rangle = \frac{\int dx\, \phi^*(x) A(\frac{\hbar}{i}\partial_x, x, t)\phi(x)}{\int dx\, \phi^*(x)\phi(x)} \tag{3.37}$$

We have used here a form that does not require explicit normalization.

3.8 Postulates of Quantum mechanics

In this section, we have gone through the process of constructing a theory, that describes new experimental facts from the double-slit experiment. We can now summarize our choices in a set of postulates.[4]

1. State: The state of a quantum mechanical system with one particle is fully described by the wave function $\Psi(\vec{r}, t)$, that depends on the position of the particle and the time. For systems with several particles, the wave function depends on the position s of all particles, i.e. $\Psi(\vec{r}_1, \ldots, \vec{r}_N, t)$.
2. Dynamics: The time evolution of a wave function is determined by the Schrödinger equation Eq. 3.29.
3. Observable: Observables are represented by linear differential operators acting on the wave function. We will see later that these differential operators need to have the property of being **hermitian**, a term that will be described later.
4. Measurement: For a given state of the system only the probabilities for the outcome of an experiment can be predicted. The mean value for the of the predicted outcomes is the expectation value defined by Eq. 3.37. In particular, the expectation value of the observable A that is defined by $A(\vec{r}, \vec{p}) = \delta(\vec{r} - \vec{r}_0)$ provides the probability density for finding the particle at position $\vec{r}_0$. This probability density is the square of the wave function at position $\vec{r}_0$, that is $\psi^*(\vec{r}, t)\psi(\vec{r}, t)$.

This set of postulates is not formulated in its most general form. This set is also not complete and a few postulates for the measurement process will be added later.

The connection with classical mechanics is given by the following two requirements.

1. if the classical dynamics is governed by Hamilton equations with a Hamiltonian $H(p, q)$, so that

   $$E = H(\vec{p}, \vec{x})$$

   the wave function evolves according the differential equation

   $$i\hbar\partial_t \Psi(\vec{r}, t) = H(\frac{\hbar}{i}\vec{\nabla}, \vec{r})\Psi(\vec{r}, t)$$

   The connection is provided by the correspondence principle Eq. 3.28.

2. If a classical observable is represented by a function $A(\vec{p}, \vec{x})$, then the quantum mechanical operator representing the observable is obtained by the correspondence principle.

3. The transformation of a function into an operator via the correspondence principle is not always unique. If the function represents a measurable quantity, one has to choose that operator, which is hermitian. Only this choice ensures that the theory predicts only real-valued measurable data.

3.9 Is Quantum theory a complete theory?

There is a fundamental question related to measurements, that is heavily discussed until today. Let us go back to the diffraction pattern of photons. A theory based on waves, namely Maxwell's equations can describe the intensity of the electromagnetic field. However, the experiment also shows that when the intensity is turned down, we observe a discrete signal here or there, but no more a continuous signal.

There are two conclusions following out of this observation.

- The events come in chunks.

- The intensity corresponds to the probability of an event.

The second observation is the reason for calling the wave function the **probability amplitude**: $\phi^*(x)\phi(x)$ is the probability to find the particle at position x.

The probabilistic nature of quantum mechanical predictions appears unsatisfactory, and seems to indicate that the theory is incomplete. The sentence "Gott würfelt nicht" ("God does not play dice") was Einstein's comment to quantum theory, but all attempts to find an inconsistency in the theory, or any contradiction between the predictions of the theory and experimental observation failed.

It is very difficult to imagine that any deterministic theory is capable of describing the outcome of the double-slit experiment with an individual particle. The experimental setup we described for the double-slit experiment in figure 2.5 is completely symmetric with respect to a mirror reflection at the $y = 0$ line connecting the source with the midpoint of the two slits. The outcome of the individual signals, however, is not symmetric. A single particle can appear anywhere, as long as the probability is finite, which in general is off-center and therefore not symmetric, unlike the experimental setup. Hence we cannot use a symmetric theory to predict the outcome for a single particle. We need a mechanism to break the symmetry. But where do we take the information **how** the symmetry is broken? Until we find some uncontrolled environmental effects not considered in the experimental setup as we described it, we need to be content with the probabilistic theory.

Do we have to accept a theory that is less complete and predictive than classical mechanics? Not really: think of chaos. If we would know the exact values of positions and velocities of a particle and the equations of motion, classical mechanics can, in principle, predict the outcome for all future. But is the exact specification of the initial conditions meaningful? Try to devise an experiment so that the trajectory is prepared in a unique way far into the future. It does not suffice to determine the initial conditions very, very accurate. Accurate is not what we mean with exact. After some time, trajectories in a chaotic system with even slightly different initial conditions deviate completely from each other. Hence, also in classical mechanics there are experiments, whose outcome we cannot predict with arbitrary certainty.

On the other hand, there are experiments for which quantum mechanics provides exact predictions. Namely those for which the probability distribution shrinks to a point.

When we compare quantum mechanics and classical mechanics closely, the difference in the completeness appears more gradual than qualitative.

3.10 Planck constant

We still need to determine the value of one undetermined constant, namely the **Planck constant** h. In practice, the **reduced Planck constant** $\hbar \stackrel{\text{def}}{=} \frac{h}{2\pi}$ is used[24] and often both are named in the same way, namely as Planck constant. Sofar the Planck constant played the role of an arbitrary proportionality constant in the correspondence principle $E = \hbar\omega$, $p = \hbar k$.

Only experiment can tell us the value of $\hbar$. The best value of $\hbar$ is currently[19, 20]

$$\hbar = 1.05457168 \times 10^{-34} \text{kg} \frac{\text{m}^2}{\text{s}} \tag{3.38}$$

What is the physical meaning of $\hbar$?

- The unit of $\hbar$ is that of an angular momentum $\vec{L} = \vec{r} \times \vec{p}$ or of a phase-space-volume element $dp\,dx$.
- Planck constant determines the length scale on which the wavy nature becomes evident, namely the wave length λ. Given a classical particle with velocity v, the wave length is $\lambda = \frac{2\pi}{k} = \frac{2\pi\hbar}{p} = \frac{2\pi\hbar}{mv}$. We recognize that the wavelength vanishes if Planck constant would be zero. In this case, we would return to classical physics, where the oscillations are so rapid that they cannot be observed.
- The limit of fast particles is similar to the classical limit $\hbar \to 0$. Indeed, quantum mechanical effects seem to become unimportant at high temperatures.
- Let us consider a few real world examples:
  - What is the wave length of a walking person? Let us make an estimate: the person should weigh 72 kg and walk with a speed of 3.6 km/h=1m/sec. Its wave length is therefore $\lambda \approx 10^{-35}$ m. This is the 10^{25}th fraction of an interatomic distance. Thus we can safely describe a person by classical physics.
  - The situation is different for an electron. An electron has a mass of about 10^{-30} kg. The wave length of an electron at the same speed is already three millimeter or about 10^7 atomic distances.

3.11 Wavy waves are chunky waves: second quantization

We have mentioned that we have an intensity which we identify with a particle number. However, sofar we have no way to explain why the intensity comes in chunks, as we expect it from particles. Here I will give a brief outlook on how the chunks are explained. I cannot give here the theory in its full glory, because that fills another lecture series on its own. Rather I want to sketch the main lines of thought.

Our starting point was the linear chain. We built it up from particles connected by springs. A modification of the linear chain revealed, that the collective motion of the beads produces wave packets, that behave like classical particles. Hence we identified particles with wave packets. To be consistent, we have to describe also every bead of the chain by a wave packet. What are the consequences?

Let us consider a violin chord as a realization of the continuous linear chain.[25] The violin chord

[24] The Planck constant is defined as h. This has been a somewhat unfortunate choice, because h appears nearly always in combination with $1/(2\pi)$. Therefore, the new symbol $\hbar = h/(2\pi)$, denoted reduced Planck constant has been introduced. The value of $\hbar$ is $\approx 10^{-34}$ Js to within 6%.

[25] The linear chain is not exactly the correct model for a violin chord. The transversal oscillation of the violin chord differs from the longitudinal oscillation of the linear chain. However, for small displacements the transversal oscillation behaves mathematically similar to a longitudinal oscillation. A better realization of the linear chain would be a rubber chord stretched along its direction.

produces waves that are enclosed by its end points. A violin chord oscillate with frequencies that are multiples of the pitch. It does not oscillate with any other frequencies.

What the violin chord is for a massless particle, the harmonic oscillator is for a massive particle. Remember, that we identified the linear chain with a massless particle. The wave function of a particle with mass in a harmonic potential oscillates with frequencies that are multiples of the classical eigenfrequency (plus a constant). The frequency is related to the energy. Thus the energy of such a system can change in discrete steps.

Let us consider a chain where we freeze all beads except one. This bead can be described by a harmonic oscillator. Its energy is discretized, like the frequency of a violin chord. Thus we have a special wave, which comes in chunks. Each frequency corresponds to a given number of particles. Since the energy of a wave is always quadratic in the wave function amplitudes, it can always be related to harmonic oscillators and all harmonic oscillators have discrete and equidistant energy levels. Hence, when the wave function amplitude (or the position of the beads) is described by wave packets, all field theories of the type discussed so far result in chunky waves.

Thus we have an additional dimension in addition to space and time, namely the field of a given particle. This system is described by a wave function $\phi(x,t)$. If we quantize this world, we obtain chunky fields and particle waves.

3.12 Summary

In this section we introduced the Schrödinger equation. In three dimensions the Schrödinger equation has the form:

$$i\hbar\partial_t\psi(\vec{r},t) = \mathcal{H}(\frac{\hbar}{i}\vec{\nabla},\vec{r})\psi(\vec{r},t) \tag{3.39}$$

It has been obtained from the classical energy expression $E = \mathcal{H}(\vec{p},\vec{r})$ which is transformed into a differential equation by using the correspondence principle, $E \to i\hbar\partial_t$ and $\vec{p} \to \frac{\hbar}{i}\vec{\nabla}$, and multiplication by the wave function.

- $\psi(\vec{r},t)$ is the **wave function**. ψ is a complex function and the absolute square value $P(\vec{r},t) = \psi^*(\vec{r},t)\psi(\vec{r},t)$ is the probability density to find the particle at point $\vec{r}$ in space at time t.
- the **Hamilton operator** $\mathcal{H}(\frac{\hbar}{i}\vec{\nabla},\vec{r})$ is obtained from the Hamilton $\mathcal{H}(\vec{p},\vec{r})$ function by replacing the momentum $\vec{p}$ by the **momentum operator**

$$\hat{\vec{p}} = \frac{\hbar}{i}\vec{\nabla} \tag{3.40}$$

- $\hbar = 1.054 \times 10^{-34}$ Js is **Planck's constant** (germ: Planck'sches Wirkungsquantum). Note that $h = 2\pi\hbar$ is called by the same name. If Planck's constant were zero, we would obtain our traditional classical point mechanics.

Instead of describing a particle as a point with defined position and a momentum, in quantum mechanics we describe particles by a wave packet with a intensity distribution $\phi^*(x)\phi(x)$ and a momentum distribution.

The wavy nature of the wave packets becomes evident only at small length scales, so that it evaded observation for a long time. One exception is again electromagnetism. There the wave length can be very large. Here a field theory has been used a long time before quantum mechanics has been invented.

We have also sketched how the quantization of the wave function itself naturally leads to the concepts of chunks of the wave, which are identified as individual particles. Thus we changed our perspective radically, from point-like particles to "waves of waves". The macroscopic predictions of quantum theory are, however, still in accord with the particle picture classical mechanics.

Chapter 4

Quantum effects

In this section, we will explore the physical effects that result from the Schrödinger equation:

QUANTUM EFFECTS

- Spreading of wave packets,
- Splitting of wave packets,
- Tunnel effect,
- Quantized energy levels,
- Zero-point energy, and
- Resonances.

Fig. 4.1: Max Planck, 1858-1947. German Physicist. Nobel price 1918 for the discovery of energy quanta. Photograph from year 1901.

One of the main tasks of this section is to become familiar with the similarities and differences of quantum-mechanical particle waves on one side and either classical particles or classical waves on the other side. The comparison will lead us to conceptual difficulties that we have to become familiar with. The reason we try to compare to classical problems is that this allows us to use the experience that we have already gained. However, it is important to also gain a thorough understanding of the conceptual problems.

In order to illustrate the quantum effects, I am choosing a class of very simple problems that can be solved analytically, namely systems where the potential varies only in one dimension and is piecewise constant. Thus the particles behave like free particles except where the value of the potential changes abruptly to another value. The problems are

- the free particle,
- the potential step,
- the particle in a box,
- the potential barrier, and
- the potential well.

4.1 Quasi-one-dimensional problems in the real world

Even though the examples are chosen to introduce quantum effects, the one-dimensional problems discussed in the following play an important role in practical problems. All these problems can be realized and have been used in electronic devices.

Electronic devices are build by bringing different materials into contact with each other, in order to control the flow of electrons through these devices. Three basic types of materials play a role: **metals**, which conduct electrons freely, **insulators** which do not conduct electrons, and **semiconductors** which can be made conducting or insulating depending on the electrostatic fields, temperature and doping.

Each material can be considered to be homogeneous and each material has its own dispersion relation for the electrons. This dispersion relation is called **band structure**[1]. Every state can be occupied by a maximum of two electrons. The **bands**, the branches in the dispersion relation, are filled with electrons up to an energy, the so-called **Fermi energy** E_F. In Fig. 4.2 we divide materials into two classes: If a band crosses the Fermi energy the material is a metal. Otherwise it is an insulator or a semiconductor. In this case, a forbidden energy region, where electrons cannot travel, separate the filled states, called the **valence bands**[2], from the empty states, called the **conduction bands**. The forbidden energy region is called the **band gap**.

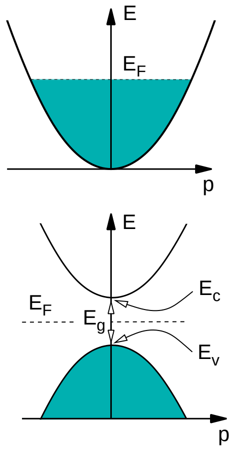


Fig. 4.2: Schematic band structure of a metal (top) and an insulator (bottom). E_F is the Fermi level. E_g is the band gap of the insulator. E_v is the top of the valence band and E_c is the bottom of the conduction band.

If there is no electric field, there are as many electrons moving right as there are electrons moving left. As a consequence there is no net electric current. An electrical current flows only when the balance of electrons moving into opposite directions is broken, which can only happen when electrons are excited above the Fermi energy. In a metal there is no minimum energy to excite electrons. Therefore, metals conduct electrons at all temperatures. In an insulator, the band gap is so large that thermal energies are insufficient to lift electrons from the valence band into the conduction band. Therefore, hardly any current flows in an insulator. In a semiconductor, the band gap is sufficiently small, so that the conductance can be easily modulated by temperature, dopant atoms and external fields.

We can learn a lot about electrons in materials from the free particle model, which underlies much of the discussion in this chapter. In a semiconductor, only the lowest part of the conduction band and the uppermost part of the valence band matter for electron transport, because thermal energies are small and cannot lift many electrons much higher. Therefore, we can perform a Taylor expansion of the conduction band about the minimum up to second order. This gives a parabola shaped dispersion relation, which can be written as

$$E(k) = E_c + \frac{\hbar^2 \vec{k}^2}{2m_c^*}$$

E_c is the conduction band minimum and m_c^* is called the **effective mass** of the conduction electron. Similarly we can describe the hole by a dispersion relation

$$E(k) = E_v - \frac{\hbar^2 \vec{k}^2}{2m_v^*}$$

where E_v is the valence band top and m_v^* is the effective mass of the holes. Holes behave as particles with a positive charge, similar to positrons.

[1] The concept of band structures is more general. There are band structures for phonons, that are the quanta of lattice vibrations. There are even band structures for photons in so-called photonic crystals, a recent and highly interesting research area.

[2] Usually one distinguishes the valence states from so-called core states, which are also occupied. Core states are located so close to the nucleus, that they do not interact with neighboring atoms.

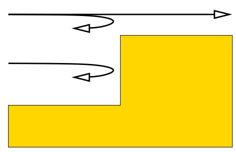
We have seen that particles which interact with a material behave like free particles but with changed properties such as the mass or the charge. Such particles are called **quasiparticles**. The situation is similar to that of light entering a material. Inside the material the speed of light is lower than in vacuum, which explains the refractive index of materials. Also here we consider the photon and the response of the material to the electric fields of the photon as one quasi particle.

We have done some simplifications: In reality (1) the effective mass can be a 3×3 tensor. (2) The minimum need not to be located at $k = 0$. (3) There may be several minima and several nearby branches of the bands at the conduction and valence band edges.

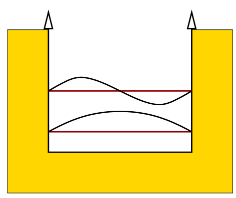
In a metal, the band structure can be described by a parabola with a minimum below the Fermi energy. Only electrons near the Fermi energy contribute to the current. Close to the Fermi energy, the dispersion relation is approximately linear with the wave vector $\vec{k}$, so that the electrons have a constant velocity similar to light and sound waves.

Here I will describe the realization of the model systems discussed in this section in electronic devices:

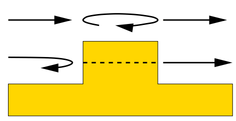
- The **free particle** describes for example the transport in a bulk material
- The **potential step** is a model for an interface between two materials with different band edges.



- In the **particle in the box** a particle is confined in all three spatial directions as in a quantum dot. A quantum dot could be realized by surrounding a semiconductor by an insulator, so that any conduction electrons cannot escape. Quantum dots are used for example for so-called single electron transistors. If the particle is confined only in one direction, the particle in the box describes a two-dimensional electron gas, which was important to obtain the so-called quantum Hall effect, for which v. Klitzing was awarded the Nobel prize.

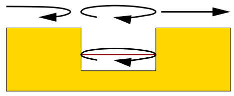


- The **finite barrier** is realized by an insulating layer between two metals or semiconductors. The finite barrier is a model of a gate oxide in a MOSFET[3] transistor to describe its tunnel current. Also the scanning tunneling microscope (Nobel prize H. Rohrer and G. Binnig) uses the tunneling current from a surface to a tip to monitor surfaces with atomic resolution.



- A **particle in a well** describes electrons in a structure that can be made by depositing two different semiconductors A and B in the sequence ABA. Since the conduction band edge of the semiconductor B is higher or lower than that of A, the conduction band edge forms either a finite barrier or a finite well. The transport through a finite well is relevant for the transport through a "resonant tunnel diode"

[3]MOSFET=Metal-oxide-semiconductor field-effect transistor



- **Dirac atom** (not yet included) The Dirac atom is the most simple one-dimensional model for an atom. It is a model for a two-dimensional electron gas or a one-dimensional conductor.

Almost any of these quasi-one-dimensional problems can be realized by growing thin layers of semiconductors onto other semiconductors. Such structures can be very complex and are called multilayer structures.

4.2 Time-independent Schrödinger equation

Before we enter the discussion of quantum effects, we need to become familiar with the time-independent Schrödinger equation. If the Hamiltonian is time independent, the Schrödinger equation can be substantially simplified. Time-independent problems are so abundant, so that most of our discussion will address those problems.

For time-independent problems, the wave function can be broken up into a product of a time-dependent part, and a time independent part, which depends only on the atomic position. The time-dependent part is trivial. The time-independent part of the wave function obeys the **time-independent Schrödinger equation**.

Let us Fourier transform the Schrödinger equation with respect to the time.[4] The time-dependent wave function $\psi(r,t)$ is related to its Fourier-transformed counterpart $\psi(r,\hbar\omega)$ by[5]

$$\psi(\vec{r},t)=\int\frac{d\omega}{2\pi}e^{-i\omega t}\psi(\vec{r},\hbar\omega) \tag{4.1}$$

Inserting the Fourier transform into the Schrödinger equation Eq. 3.29 yields[6]

$$\begin{aligned}
&\left(i\hbar\partial_t-H(\frac{\hbar}{i}\vec{\nabla},\vec{r})\right)\Psi(r,t)=0\\
\stackrel{\text{Eq. 4.1}}{\Rightarrow}\quad&\left(i\hbar\partial_t-H(\frac{\hbar}{i}\vec{\nabla},\vec{r})\right)\left[\int\frac{d\omega}{2\pi}e^{-i\omega t}\psi(\vec{r},\hbar\omega)\right]=0\\
\Rightarrow\quad&\left[\int\frac{d\omega}{2\pi}\left(i\hbar\partial_t-H(\frac{\hbar}{i}\vec{\nabla},\vec{r})\right)e^{-i\omega t}\psi(\vec{r},\hbar\omega)\right]=0\\
\Rightarrow\quad&\int\frac{d\omega}{2\pi}e^{-i\omega t}\underbrace{\left[\left(\hbar\omega-H(\frac{\hbar}{i}\vec{\nabla},\vec{r})\right)\psi(\vec{r},\hbar\omega)\right]}_{\text{equation holds only if this vanishes}}=0
\end{aligned}$$

[4] Note here that we denote the Fourier transform with the same symbol as the original functions. Mathematically this is incorrect! However, this notation is very convenient when used cautiously. When we evaluate the function value we need not only the numeric value of the argument, but also the type. So if we evaluate the function for a given time, we obtain a different result than if we evaluate the function for a frequency with the same numeric value. We will later see how such a notation can be formulated rigorously in the bracket formalism of Dirac.

[5] The reader may be wonder why we use $e^{-i\omega t}$ instead of $e^{i\omega t}$, which would be our regular notation of a Fourier transform. The origin of te minus sign is that we derived the Schrödinger equation from plane waves $e^{i(kx-\omega t)}$. The sign convention for ω is a matter of choice, but if we use this form of the Schrödinger equation, we need to adhere to this choice.

[6] Plane waves $e^{i\omega t}$ are linear independent. Linear independency implies that no superposition $\sum_i c_i e^{i\omega_i t}$ of different $(\omega_i\neq\omega_j)$ plane waves can result in a vanishing function. Thus the expression

$$\int\frac{d\omega}{2\pi}c(\omega)e^{i\omega t}=0$$

can hold for all t only, when $c(\omega)=0$. Because plane waves are linear independent, the Fourier transform is unique.

We obtain

$$H(\frac{\hbar}{i}\vec{\nabla},\vec{r})\psi(\vec{r},\hbar\omega)=\psi(\vec{r},\hbar\omega)\hbar\omega$$

after realizing that the equation must be valid for every Fourier component.

By replacing $\hbar\omega$ with E, we obtain the time-independent Schrödinger equation.

TIME-INDEPENDENT SCHRÖDINGER EQUATION

$$H(\vec{r},\frac{\hbar}{i}\vec{\nabla})\ \psi(\vec{r},E)=\psi(\vec{r},E)\ E \tag{4.2}$$

We call such an equation an **eigenvalue equation**. The equation has solutions only for specific values of E. These values are the **eigenvalues** of the Hamiltonian. The set of all eigenvalues is called the **spectrum**. The solution $\psi(\vec{r},E)$ is called an **eigenstate** of the Hamiltonian.

An eigenvalue equation is similar to the problem of diagonalizing a matrix, $\boldsymbol{A}\vec{c}_n=\vec{c}_n\lambda_n$. The only difference is that the differential operator is replaced by a matrix and the vector is replaced by a function in space. We will see later that this difference is smaller than one might think. The difference is related only to the fact that in one case we describe finite matrices and vectors and in the other case infinite vectors and matrices. We will return to this aspect in detail in chapter 5.

There may be several solutions to a given energy. We call such solutions **degenerate**. The time-independent Schrödinger equation may have solutions for discrete values for the energy E, or for continuous sections of the energy axis. The set of allowed energy values is called the **eigenvalue spectrum** of the Hamilton operator. The energy spectrum of the Hamilton operator is bound from below, which is exploited in the variational principle described in Section 13.5.

Once we know all eigenstates $\psi_\sigma(\vec{r},E)$, of the time-independent Schrödinger equation, we can determine the wave function for all times as

$$\psi(\vec{r},t)\stackrel{\text{Eq. 4.1}}{=}\int\frac{dE}{2\pi\hbar}\ e^{-iEt/\hbar}\sum_\sigma\psi_\sigma(\vec{r},E)a_\sigma(E)$$

We use the index σ to distinguish different eigenfunctions for the same eigenvalue.

The initial conditions, represented by the values $\psi(r,t=0)$, determine the coefficients $a_\sigma(E)$ of this expansion in stationary states.

$$\psi(r,t=0)=\int\frac{dE}{2\pi\hbar}\sum_\sigma\psi_\sigma(r,E)a_\sigma(E)$$

4.3 Method of separation of variables

In this chapter we will work almost exclusively with one-dimensional problems. One may ask how much this has to do with the real three-dimensional world? The answer is given by the **method of separation of variables**.

As in classical mechanics, it is often possible to exploit symmetries to break down a problem into problems of lower dimensions, and, if we are lucky, into one-dimensional problems. An example is a problem with a potential that varies only in one dimension, so that it is translationally invariant in the other directions.

Quasi-onedimensional problem in three dimensions

Let us try to exploit this symmetry on a quasi-onedimensional problem: consider the Schrödinger equation with a potential that varies only in x-direction:

$$i\hbar\partial_t\Psi(\vec{r},t)=\Big[\frac{-\hbar^2}{2m}\nabla^2+V(x)\Big]\Psi(\vec{r},t) \tag{4.3}$$

We can write the Hamiltonian Eq. 4.3 as sum $H = H_1 + H_2 + H_3$ of three terms, where each acts on one variable

$$H_1 = \frac{-\hbar^2}{2m}\partial_x^2 + V(x); \qquad H_2 = \frac{-\hbar^2}{2m}\partial_y^2; \qquad H_3 = \frac{-\hbar^2}{2m}\partial_z^2$$

We make a product ansatz for the wave function

$$\Psi(\vec{r}, t) = \phi_1(x, t)\phi_2(y, t)\phi_3(z, t) \tag{4.4}$$

where each term ϕ_i solves the corresponding one-dimensional Schrödinger equation.

$$\begin{aligned}
i\hbar\partial_t\phi_1(x, t) &= \left[\frac{-\hbar^2}{2m}\partial_x^2 + V(x)\right]\phi_1(x, t)\\
i\hbar\partial_t\phi_2(y, t) &= \frac{-\hbar^2}{2m}\partial_y^2\phi_2(y, t)\\
i\hbar\partial_t\phi_3(z, t) &= \frac{-\hbar^2}{2m}\partial_z^2\phi_3(z, t)
\end{aligned} \tag{4.5}$$

In order to show that this ansatz yields the correct result, we insert it into the Schrödinger equation Eq. 4.3. We divide the equation by the wave function, which allows us to write the equations in a more economical fashion.

$$\begin{aligned}
\frac{1}{\Psi(\vec{r}, t)} i\hbar\partial_t\Psi(\vec{r}, t) &\overset{\text{product rule}}{=} \frac{i\hbar\partial_t\phi_1(x, t)}{\phi_1(x, t)} + \frac{i\hbar\partial_t\phi_2(y, t)}{\phi_2(y, t)} + \frac{i\hbar\partial_t\phi_3(z, t)}{\phi_3(z, t)}\\
&\overset{\text{Eq. 4.5}}{=} \frac{1}{\phi_1(x, t)}\left[\frac{-\hbar^2}{2m}\partial_x^2 + V(x)\right]\phi_1(x, t)\\
&\quad + \frac{1}{\phi_2(y, t)}\left[\frac{-\hbar^2}{2m}\partial_y^2\right]\phi_2(y, t) + \frac{1}{\phi_3(z, t)}\left[\frac{-\hbar^2}{2m}\partial_z^2\right]\phi_3(z, t)
\end{aligned}$$

Now there is the only tricky part: We can multiply the first term on the right-hand side with $\phi_2\phi_3$ from the right side and divide by the same on the left hand side. This does not change the result, because the Hamiltonian H_1 does not act on $\phi_2\phi_3$. Similarly, we multiply the second term with $\phi_1\phi_3$ on the right and divide by the same factor on the left. Finally, we multiply the second term with $\phi_1\phi_2$ on the right and divide by the same factor on the left.

Thus we obtain

$$\begin{aligned}
\frac{1}{\Psi(\vec{r}, t)} i\hbar\partial_t\Psi(\vec{r}, t) &= \frac{1}{\Psi(\vec{r}, t)}\left[\frac{-\hbar^2}{2m}\partial_x^2 + V(x)\right]\Psi(\vec{r}, t)\\
&\quad + \frac{1}{\Psi(\vec{r}, t)}\left[\frac{-\hbar^2}{2m}\partial_y^2\right]\Psi(\vec{r}, t) + \frac{1}{\Psi(\vec{r}, t)}\left[\frac{-\hbar^2}{2m}\partial_z^2\right]\Psi(\vec{r}, t)\\
\Rightarrow i\hbar\partial_t\Psi(\vec{r}, t) &= \left[\frac{-\hbar^2}{2m}\partial_x^2 + V(x) + \frac{-\hbar^2}{2m}\partial_y^2 + \frac{-\hbar^2}{2m}\partial_z^2\right]\Psi(\vec{r}, t) = H\Psi(\vec{r}, t)
\end{aligned}$$

This proves that our ansatz, Eq. 4.4 with Eq. 4.5, yields the correct solution.

General case

The method of separation of variables is not limited to quasi-one-dimensional problems. It can be applied to a whole class of Hamiltonians. Let us state it in some greater generality.

METHOD OF SEPARATION OF VARIABLES

If the Hamiltonian is a sum of several parts, where each part only acts on one variable such as

$$i\hbar\partial_t\Psi(x_1, x_2, \dots, t) = \left[\sum_j H_j(\frac{\hbar}{i}\partial_j, x_j, t)\right]\Psi(x_1, x_2, \dots, t) \tag{4.6}$$

then the wave function can be expressed as

$$\Psi(x_1, x_2, \dots, t) = \prod_j \phi_j(x_i, t) \tag{4.7}$$

with

$$i\hbar\partial_t\phi_j(x, t) = H_j(\frac{\hbar}{i}\partial_j, x, t)\phi_j(x, t) \tag{4.8}$$

Thus we have broken down the multidimensional problem into many one-dimensional problems, that are much easier to solve. Of course, due to the superposition principle, any superposition of such solutions is again a solution.

In the proof, we exploit that the Hamilton operator $H_i(\frac{\hbar}{i}\partial_i, x_i, t)$ does not act on wave functions $\phi_j(x_j, t)$ with $j \neq i$. They do not depend on the parameter x_i on which the Hamilton operator $H_i(\frac{\hbar}{i}\partial_i, x_i, t)$ acts. Therefore, $H_i(\frac{\hbar}{i}\partial_i, x_i, t)\phi_j(x_j, t) = \phi_j(x_j, t)H_i(\frac{\hbar}{i}\partial_i, x_i, t)$ for $j \neq i$.

The proof of Eq. 4.6 with Eqs. 4.7,4.8 goes as follows

$$\begin{aligned}
\underbrace{i\hbar\partial_t\prod_i \phi_i(x_i, t)}_{\Psi(\vec{r},t)} &\stackrel{\text{product rule}}{=} [i\hbar\partial_t\phi_1(x_i, t)]\,\phi_2(x_2, t)\cdots + \phi_1(x_1, t)\,[i\hbar\partial_t\phi_2(x_2, t)]\cdots + \dots \\
&= \left[\prod_i \phi_i(x_i, t)\right]\sum_i \frac{1}{\phi_i(x_i, t)} i\hbar\partial_t\phi_i(x_i, t) \\
&\stackrel{\text{Eq. 4.8}}{=} \left[\prod_i \phi_i(x_i, t)\right]\sum_i \frac{1}{\phi_i(x_i, t)} H_i(\frac{\hbar}{i}\partial_i, x_i, t)\phi_i(x_i, t) \\
&= \sum_i H_i(\frac{\hbar}{i}\partial_i, x_i, t)\underbrace{\left[\prod_i \phi_i(x_i, t)\right]}_{\Psi(\vec{r},t)} \qquad q.e.d.
\end{aligned}$$

The method of separation of variables can be used, if we are able to express the Hamilton operator by a sum of differential operators that only depend on a single variable. Sometimes the choice is not immediately apparent and a coordinate transform is required to bring it into the desired form. The methods of finding such variable transforms are similar to the techniques used in classical mechanics, such as the transformation on center-of-mass and relative coordinates in a many particle system and the division into angular and radial coordinates in a spherical system.

Most Hamiltonians, however, cannot be written in this simple form. Even if the problem cannot be separated exactly, the method of separation of variables is invaluable in finding approximate solutions. As most quantum mechanical problems cannot be solved analytically or even numerically, it is important to find models that represent the true system closely – even though not exactly –, but contain symmetries that allow to break down the problem. This approximate model is the starting point for approximation techniques discussed in section 13 to include whatever has been neglected in the model system.

4.4 Free particles and spreading wave packets

A free particle does not experience forces. It experiences a constant potential. Thus, it obeys the Schrödinger equation

$$\left[\frac{-\hbar^2}{2m}\partial_x^2+V_0\right]\psi(x,E)=E\psi(x,E) \tag{4.9}$$

with a spatially constant potential $V(x)=V_0$.

The solutions can be found using the Ansatz $\psi(x,E)=\mathrm{e}^{ikx}$. By insertion we find the dispersion relation

$$E(k)=\frac{\hbar^2k^2}{2m}+V_0$$

The energy eigenstates of a free particle are plane waves

$$\Psi(x,E)=\mathrm{e}^{ik(E)x} \tag{4.10}$$

with a wave vector k that has length $k(E)=\pm\frac{1}{\hbar}\sqrt{2m(E-V_0)}$.

We have discarded the solutions with $E<V_0$. They are not plane waves, but exponential functions. Neither plane waves nor exponential functions are normalizable. Therefore, they cannot describe the wave function of a single particle. While for the plane waves, a way out of this dilemma has been found, for the exponential function the problem cannot be cured. Nevertheless, exponential solutions will play a role as partial solution in problems that have been pieced together.

A few properties are noteworthy:

- The intensity $\phi^*(x)\phi(x)=\mathrm{e}^{-ikx}\mathrm{e}^{ikx}=1$ of the plane wave is uniformly distributed over all space.
- The momentum of a plane wave is $\langle p\rangle=\frac{\int dx\psi^*(x)\frac{\hbar}{i}\partial_x\psi(x)}{\int dx\psi^*(r)\psi(r)}=\hbar k$. Furthermore, the spread of momenta $\langle(p-\langle p\rangle)^2\rangle=\langle p^2\rangle-\langle p\rangle^2$ vanishes for a plane wave.

Thus a plane wave describes a beam of particles flying with the same momentum p, or in other words, constant velocity $v=\partial_pE(p)=\frac{p}{m}$.

Free particle in three dimensions

We can use the result for the free particle in one dimension and the method of separation of variables to obtain a solution for a free particle in three dimensions: The wave function in three dimensions is the product of the three one-dimensional free particles.

$$\Psi(\vec{r})=\mathrm{e}^{ik_xx}\mathrm{e}^{ik_yy}\mathrm{e}^{ik_zz}=\mathrm{e}^{i\vec{k}\vec{r}}$$

Here it becomes apparent why we denote $\vec{k}$ as wave vector.

The energy is the sum of the three one-dimensional problems

$$E(\vec{k})=\frac{(\hbar\vec{k})^2}{2m}+V_0$$

and the momentum is

$$\langle\vec{p}\rangle\overset{Eqs.\ 3.28,3.37}{=}\frac{\int d^3r\ \Psi^*(\vec{r})\frac{\hbar}{i}\vec{\nabla}\Psi(\vec{r})}{\int d^3r\ \Psi^*(\vec{r})\Psi(\vec{r})}=\hbar\vec{k}$$

The velocity[7] of the particles in the beam is then $\vec{v}\overset{\text{Eq. 3.6}}{=}\vec{\nabla}E(\vec{p})=\frac{1}{m}\vec{p}$.

[7] The group velocity of a wave packet has been linked to the dispersion relation $\omega(k)$. The Schrödinger equation has been constructed such that its dispersion relation is analogous to that of classical particles $E=H(p,x,t)$. Therefore, we may use Hamilton's equations, that in principle are equations of motion for classical particles.

Wave functions that are not normalizable

There is a problem related to solution Eq. 4.10: It cannot be normalized, because the integral Eq. 3.33 is infinite. This would contradict the interpretation of the wave function as probability amplitude.

There are two possible ways around the problem.

1. Instead of an infinite integration volume we consider a large but finite volume, such as the universe, which in fact is finite. In the finite volume the wave function is normalizable. Furthermore, the spectrum in a finite volume is discrete, and does not contain continuous sections. The disadvantage is that the amplitude of the wave function is extremely tiny, if the volume is very large. This point of view, which we will adopt in most of this book, avoids mathematical subtleties. Before the expressions are evaluated in practice, one has to take the limit of infinite volume.

2. We consider the non-normalizable wave function only as one part of the true wave function. The true wave function would be a product of the non-normalizable wave function with an envelope function that itself is normalizable, such as the wave packet

$$\psi(\vec{r},t)=\chi(\vec{r},t)\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}$$

   where the plane wave is not normalizable, but the envelope function $\chi(x,t)$ is normalized. By multiplication with a normalizable envelope function, the complete wave function is normalizable. The corresponding energy eigenvalue spectrum may contain discrete eigenvalues and, simultaneously, energy regions with a continuous set of eigenvalues.

The second choice is the mathematically more sophisticated approach. In a way, it limits the integration volume like the first suggestion. It does it in a way so that the wave function remains differentiable everywhere in space.

4.4.1 Wave packets delocalize with time

Let us explore, what happens with a free particle that is localized in space. We start with a wave packet of Gaussian shape[8]. We use the ansatz

$$\psi(x,t)=A(t)\mathrm{e}^{-\frac{1}{2}B(t)x^2}\ , \tag{4.11}$$

which leads to differential equations for $A(t)$ and $B(t)$. Because the derivation is a bit lengthy, we solve the problem in the App. B. Here, we only report the solution:

$$\psi(x,t)\overset{\text{Eq. }B.16}{=}\sqrt[4]{\frac{C_0}{\pi|C(t)|^2}}\mathrm{e}^{-\frac{1}{2C(t)}x^2}\qquad\text{with}\qquad C(t)\overset{\text{Eq. }B.11}{=}C_0+\frac{i\hbar t}{m} \tag{4.12}$$

The symbol C_0 is the initial value of $C(t)$. The wavefunction acquires a complex phase, because $C(t)$ becomes complex. As a result the wave function oscillates in space.

What we are mostly interested in is the probability to find the particle at a certain point in space. Let us therefore evaluate the probability distribution $P(x)=\psi^*(x,t)\psi(x,t)$.The derivation is found again in appendix B. Here we only report the result.

$$P(x)\overset{\text{Eq. }B.17}{=}\frac{1}{\sqrt{2\pi}\cdot d(t)}\mathrm{e}^{-\frac{1}{2}\left(\frac{x}{d(t)}\right)^2}\qquad\text{with}\qquad d(t)\overset{\text{Eq. }B.18}{=}d(0)\sqrt{1+\left(\frac{\hbar t}{2md^2(0)}\right)^2} \tag{4.13}$$

The initial width is related to the coefficient C in Eq. B.11 by $d(0)=\frac{1}{2}\sqrt{C_0}$. The density maintains its Gaussian shape, but the width, which is proportional to $d(t)$ grows, after a certain while, almost linearly with time, that is $d(t)\to\frac{\hbar t}{2md(0)}$ for $t\to\infty$.

[8] A Gaussian has the form e^{-ax^2}.

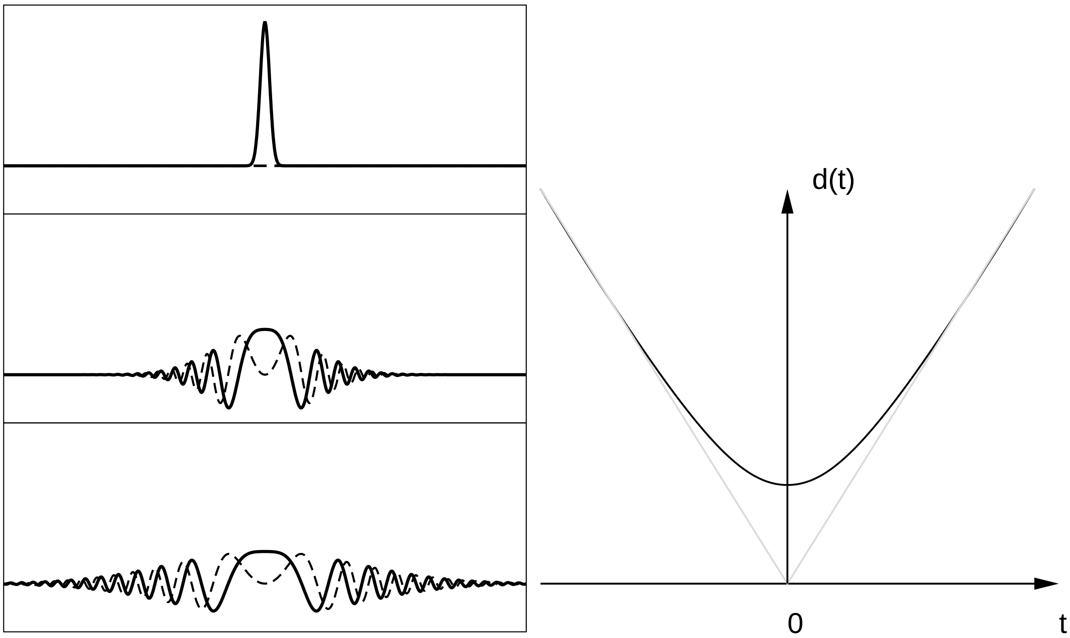


Fig. 4.3: Decay of a Gaussian wave packet with time. The horizontal axis is a spatial coordinate. The full line is the real part of the wave function and the dashed line is the imaginary part. The three figure represent the initial state and equi-distant later times. Note that the absolute square of the wave packet maintains its Gaussian shape. Real and imaginary part however oscillate. The oscillation is more rapid in the wings: The wings are made of the contributions that run away most quickly, which also have the smallest wavelength.($v_g = \frac{1}{m}\hbar k$).

Something strange happens here: the smaller the wave packet is initially, the faster it spreads out. In other words, a small width $d(0)$ leads to a rapid growth of the width as seen from Eq. 4.13. This observation can be rationalized as follows: When we build a narrow Gaussian from plane waves,

$$\frac{1}{\sqrt{2\pi}\cdot d}\mathrm{e}^{-\frac{1}{2}\left(\frac{x}{d}\right)^2} = \int \frac{dk}{2\pi}\,\mathrm{e}^{ikx}\,\mathrm{e}^{-\frac{1}{2}(d\cdot k)^2}, \tag{4.14}$$

we have to include plane waves e^{ikx} with large kinetic energy $\hbar\omega(k)$. Those contributions have high velocity and run away quickly.

What we see here, is another manifestation of **Heisenberg's uncertainty relation** , which will be derived later in section 7.4. Heisenberg's uncertainty relation says that if the position of a particle is well known, there is a large uncertainty about its momentum. In other words, if the position is precisely determined, that is if the width of the wave packet is small, the spread in velocities is large.

Delocalization in practice

Let us consider a person who has a mass of 50 kg and let us consider that its position is initially certain to within the thickness of a hair that is $d(0) = 1\,\mu$m. How long does it take until the position becomes uncertain to about $d(t) = 1$ mm. With $\hbar \approx 10^{-34}$kg m^2/s, we obtain $t \overset{\text{Eq. 4.13}}{=} \frac{1}{\hbar}2md(0)d(t) = 10^{27}$ s. Because a year has about 3×10^7 s, we have to wait 10^{20} years!

On the other hand, let us consider a microscopic particle such as an electron: The electron has a mass of about 10^{-32} kg. To spread from 1 μm to 1 mm, takes only one tenth of a microsecond.

Now let us consider a truly microscopic situation: Let us localize an electron on a single atom, which has a radius of about 0.1 nm=10^{-10} m. What is the localization region after a typical atomic oscillation period of about 10^{-13} s? Its diameter has grown to a few μm or 10,000 times the size of an atom. This indicates that the atoms do not experience the electrons as "particles" but as delocalized waves.

This discussion explains why delocalization as quantum effect is not observable in our daily life, while it is essential in the nano-world.

4.4.2 Wick rotation

The decaying wave packet reminds us of a diffusion process. While the process of a diffusing particle is entirely different from that of a spreading wave packet, there is an intrinsic connection on a mathematical level. The diffusion equation is closely related to the Schrödinger equation. Consider the diffusion equation

$$\partial_t \rho(\vec{r}, t) = D\vec{\nabla}^2 \rho(\vec{r}, t)$$

where ρ is the density and D is the isotropic diffusion constant. Now let us perform the so-called **Wick-rotation**

$$D \rightarrow \frac{\hbar^2}{2m} \qquad \text{and} \qquad t \rightarrow \frac{it}{\hbar} \qquad \text{and} \qquad \rho \rightarrow \psi$$

then we obtain from the diffusion equation the Schrödinger equation of a free particle.

$$i\hbar\partial_t \psi(\vec{r}, t) = -\frac{\hbar}{2m}\vec{\nabla}^2 \psi(\vec{r}, t)$$

The close connection between quantum mechanics and statistical mechanics, which has become evident here, has been extremely fruitful in the past and has lead to the Quantum Monte Carlo method and the Renormalization theory for the description of phase transitions (Kenneth G. Wilson, Nobel prize for Physics 1982).

4.4.3 Galilei invariance (home study)

We have seen that the wave function of a resting wave packet spreads equally in all directions. The analogy with the diffusion equations may lead to a misunderstanding that this is also true if the particle have initially a non-zero mean velocity. Such a behavior would violate Galilei invariance, saying that the laws of nature are independent of the velocity of an observer. This is the reason for investigating a wave function with a finite velocity and its relation to the resting wave packet.

Galilei invariance implies that the laws of nature are independent of a (constant) velocity of the observer. In other words, there is no absolute velocity.[9] Here we illustrate, rather than prove, Galileo invariance of the Schrödinger equation for a free particle by investigating how a wave packet changes its shape with time.

We will see that envelope function of a free particle does not remain constant with time: Firstly, it moves with a velocity $v = \frac{dE}{dp}$ and secondly it changes its shape slowly so that eventually it covers all space. We will show that a moving wave packet behaves exactly the same way as a resting one.

In order to show this equivalence between a moving wave packet and a resting one, we propose a transformation from a resting wave packet to one moving with velocity v. Then we show that the transformed wave packet obeys the Schrödinger equation if the resting does as well.

If $\Psi_0(x, t)$, the "resting wave packet", obeys the Schrödinger equation for a free particle, then also the wave function

$$\Psi_v(x, t) \stackrel{\text{def}}{=} \Psi_0(x - vt, t)e^{i(kx - \omega(k)t)} , \tag{4.15}$$

[9]Note that the problem with an external, spatially dependent, potential $V(x)$ is not Galileo invariant, because that potential is fixed in space. However, if we step back and include whatever creates that potential into the description. This enlarged system is Galileo invariant.

the "moving wave packet,does", if v is the group velocity

$$v = \partial_k \omega$$

Interesting is that the transformation does not only correspond to a time dependent shift of the wave function, but that it obtains in addition a phase factor, the plane wave. A more detailed derivation is given in appendix A on p.259.

Now we show by insertion that the moving wave packet $\Psi_v(x,t)$ as defined in Eq. 4.15 obeys the Schrödinger equation.

$$\begin{aligned}
0 &\stackrel{?}{=} \left[i\hbar\frac{d}{dt} - \frac{-\hbar^2}{2m}\partial_x^2\right]\Psi_0(\underbrace{x-vt}_{\stackrel{\text{def}}{=}y(x,t)},\underbrace{t}_{\stackrel{\text{def}}{=}\tau(x,t)})\mathrm{e}^{i[kx-\omega t]} \\
&= \mathrm{e}^{i[kx-\omega t]}\left[-i\hbar v\partial_y + i\hbar\partial_\tau + \hbar\omega + \frac{\hbar^2}{2m}(\partial_y^2 + 2ik\partial_y - k^2)\right]\Psi_0(y(x,t),\tau(x,t)) \\
&= \mathrm{e}^{i[kx-\omega t]}\left[i\hbar\partial_\tau + \frac{\hbar^2}{2m}\partial_y^2 - \underbrace{(v - \frac{\hbar k}{m})}_{=0} i\hbar\partial_y + \underbrace{(\hbar\omega - \frac{\hbar^2}{2m}k^2)}_{=0}\right]\Psi_0(y,\tau) \\
&= \mathrm{e}^{i[kx-\omega t]}\underbrace{\left[i\hbar\partial_\tau - \frac{-\hbar^2}{2m}\partial_y^2\right]\Psi_0(y,\tau)}_{=0} = 0
\end{aligned} \tag{4.16}$$

Thus, we have shown that $\Psi_v(x,t)$, as defined in Eq. 4.15, is a solution to the Schrödinger equation for a free particle, if

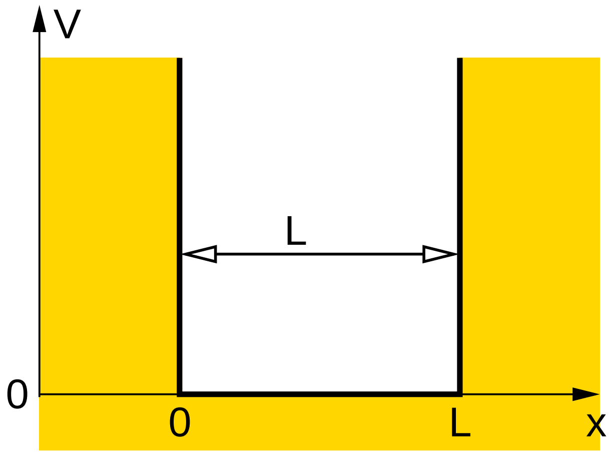
- $k = \frac{1}{\hbar}mv$,
- $\omega = \frac{1}{\hbar}\frac{\hbar^2 k^2}{2m} = \frac{1}{\hbar}\frac{1}{2}mv^2$ and, if
- Ψ_0 obeys the Schrödinger equation for a free particle as well.

Note that the total derivative d/dt acts also on the argument $y(x,t) = x - vt$, whereas the partial derivative only acts on the second argument of χ. In the last step we simplified using the dispersion relation and $v = \frac{\partial\omega}{\partial k} = \frac{\hbar k}{m}$

Eq. 4.16 is nothing but the Schrödinger equation for a particle at rest. Thus the shape of the envelope function is independent of the velocity of the wave packet. This is Galileo invariance: The equations of motion do not depend on the relative velocity between the object and the observer.

4.5 Particle in a box, quantized energies and zero-point motion

The particle in a box is our simple model for an atom. In a hydrogen atom, the electron is bound to the nucleus by electrostatic attraction. Under normal circumstances, the electron cannot overcome this attraction. We simplify the problem by replacing the Coulomb potential of the nucleus by a constant inside the attraction region and infinitely high potential walls, which ensures that the electron does not get lost.



The potential $V(x)$ for particle in a box enclosed by impenetrable walls is

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

where L is the width of the box.

Apart from the walls, the problem is similar to that of a free particle. The Schrödinger equation in the well is

$$\frac{-\hbar^2}{2m}\partial_x^2\psi(x, E) = E\psi(x, E) \qquad \text{for } x \in [0, L]$$

In the well, we use the same Ansatz as for the free particle

$$\psi(x) = \begin{cases} A\mathrm{e}^{ikx} + B\mathrm{e}^{-ikx} & \text{for } x \in [0, L] \\ 0 & \text{otherwise} \end{cases}$$

with $k(E) = \frac{1}{\hbar}\sqrt{2mE}$. Outside the well the wave function vanishes.

The potential imposes an additional boundary condition on the wave function:

$$\psi(0) = 0 \qquad \text{and} \qquad \psi(L) = 0 \tag{4.17}$$

The wave function vanishes continuously beyond the boundary of the well. At this point the motivation for these boundary conditions is based only on physical grounds, namely that the probability density in the region with infinite potential, that is outside the well must vanish. At the boundary of the box the wave function is not differentiable. This reflects that the kinetic energy in the forbidden region, and hence the curvature of the wave function, is infinite. A more rigorous justification is obtained from the solution for the particle in a box with finite walls, the particle well, which is discussed in section 4.8. We can use that solution and form the limit when the walls grow to infinity.

We satisfy the first boundary condition Eq. 4.17

$$\begin{aligned} \psi(0) = A + B = 0 \qquad &\Rightarrow \qquad B = -A \\ \Rightarrow \psi(x) = = A\left(\mathrm{e}^{ikx} - \mathrm{e}^{-ikx}\right) &= \underbrace{2iA}_{=:C}\sin(kx) = C\sin(kx) \end{aligned} \tag{4.18}$$

where we introduced a new variable $C \stackrel{\text{def}}{=} 2iA$.

The second boundary condition Eq. 4.17 translates into

$$\psi(L) = C\sin(kL) = 0 \qquad \Rightarrow \qquad kL = \pi n$$

where n is an arbitrary integer.[10]

Thus the wave vector can assume one of the discrete values k_n, where

$$k_n = \frac{\pi}{L} n \tag{4.19}$$

The dispersion relation $E = (\hbar k)^2/2m$ provides the corresponding energy eigenvalues.

$$E_n = \frac{(\hbar k_n)^2}{2m} \stackrel{\text{Eq. 4.19}}{=} \left(\frac{\hbar^2\pi^2}{2mL^2}\right) n^2 \tag{4.20}$$

where n is an integer[11]. Thus, in a box, only discrete wave vectors and eigenvalues are permitted. They are **quantized**.

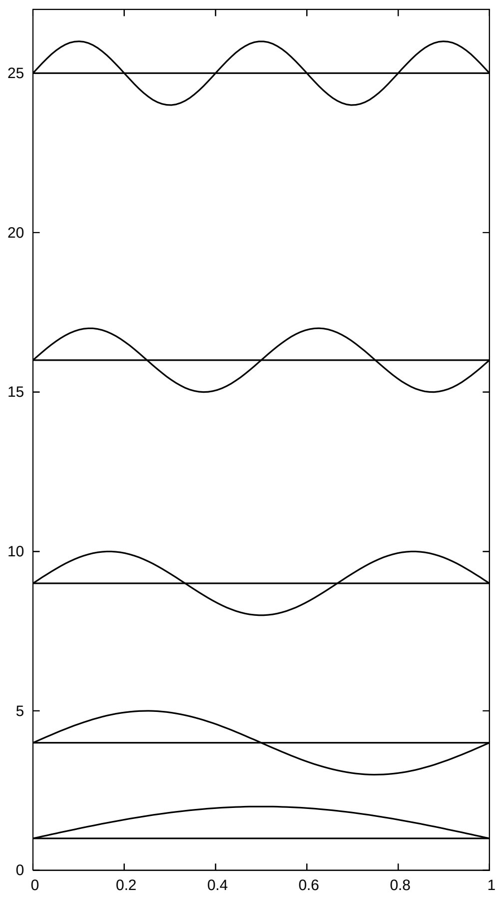


Fig. 4.4: The lowest five wave functions of the particle in the box. The vertical axis is the energy in units of $\hbar^2\pi^2/(2mL^2)$. The energy value has been taken as the zero axis for the corresponding wave function. The latter is given in arbitrary units. Note that the lowest state is not at energy zero, which is a consequence of the zero point vibration or Heisenberg's uncertainty relation. Note that the wave functions can be classified according to their behavior under mirror operation about the center of the box: Some are symmetric and some are antisymmetric. This is a consequence of the mirror symmetry of the potential. Note that the number of nodes (the zeros) increases with energy. Also this is a general feature of Problems with a local potential.

Finally we need to normalize the wave function, which determines C.

$$\int_{-\infty}^{\infty} dx\ \psi^*(x)\psi(x) = 1$$

[10] The zeros of the sinus function lie at multiples of π.

[11] integer number=ganze Zahl

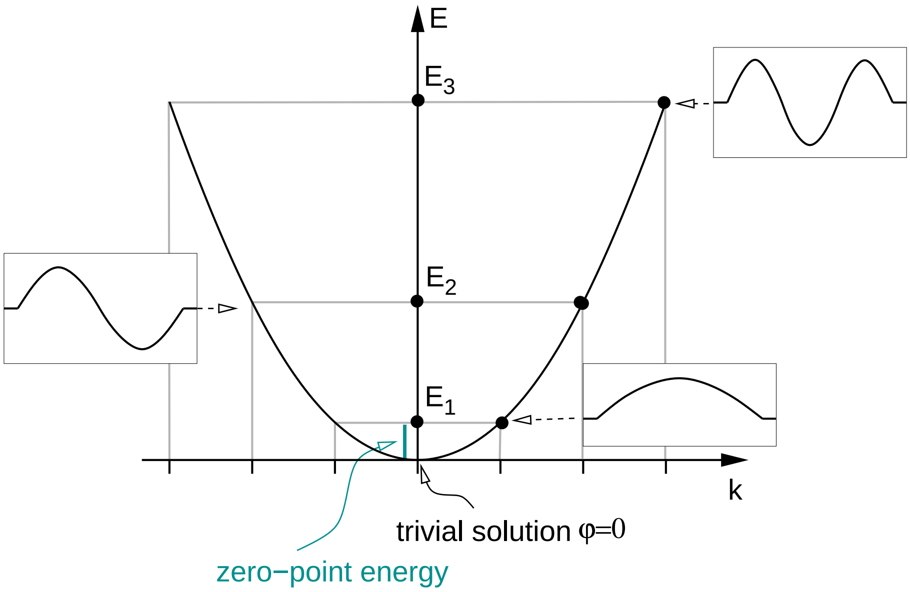


Fig. 4.5: The energy levels for a particle in a box are obtained from the dispersion relation $E(k)$ of a free particle by selecting equi-distant values $k_n = \frac{\pi}{L}\,n$ for the k-points. The spacing of the k-points changes inversely with the size of the box. The number of nodes increase with energy. Note that the wave function for $k = 0$ does not exist, because it is identical to zero. Every wave function is a standing wave composed of two plane waves with positive and negative k. They are shown in the insets.

which implies

$$
\begin{aligned}
1 &= \int_{-\infty}^{\infty} dx\; \psi^*(x)\psi(x) \stackrel{\text{Eq. 4.18}}{=} \int_0^L dx\; C^*C\sin^2(k_n x) \\
&= C^*C\int_0^L dx\; \left(\frac{1}{2}-\frac{1}{2}\cos(2k_n x)\right) = C^*C\frac{L}{2} \\
\Rightarrow |C| &= \sqrt{\frac{2}{L}}
\end{aligned}
$$

We may use C as real number.

Thus we obtain the wave functions and their energies as

$$
\begin{aligned}
\psi_n(x) &= \sqrt{\frac{2}{L}}\sin(k_n x) \qquad \text{with} \qquad k_n \stackrel{\text{Eq. 4.19}}{=} \frac{\pi}{L}n \\
E_n &\stackrel{\text{Eq. 4.20}}{=} \frac{\hbar^2\pi^2}{2mL^2}n^2
\end{aligned}
$$

We make some interesting observations:

- The lowest energy is not zero! Since the particle in the box has only kinetic energy, this implies that the particle can not be at rest.[12] This effect, which is a general feature of bound systems, is called the **zero-point motion** and the corresponding energy is called **zero-point energy**.

  We observe here an effect of **Heisenberg's uncertainty principle**, which says that the position and the momentum of a particle cannot be determined simultaneously with absolute certainty.

[12] Note that we use here the language from classical mechanics for a quantum particle.

If the particle were at rest, its momentum would be zero and therefore certain. Consequently, the position should be completely unknown, which contradicts with the fact that the particle is confined in a box. The smaller the box the larger is the zero point energy. We will later learn about a more rigorous formulation of Heisenberg's uncertainty principle.

Loosely termed, Heisenberg's uncertainty principle says that one cannot form a localized wave packet with a single plane wave. Very localized wave packets require large sets of plane waves and therefore have wide distribution of momenta. The average momentum, on the other hand is still a single quantity. In a language of particles, the uncertainty principle is most puzzling. However, in a wave description it is almost a triviality.

In a classical description an electron would fall right in the nucleus, since as accelerated charged particles it emits electromagnetic waves, and therefore looses energy. The zero point energy prevents this process.

- The allowed energies are **quantized**, that is they are discrete. This implies that we cannot add energy to the system or remove energy from the system in arbitrary amounts. The energy spacing is inversely proportional with the mass of the particle. Hence, for a very heavy particle it would not be observable and the behavior would appear classical.

  A word of caution is required here. The quantization only concerns stationary states. It is possible to prepare a state with arbitrary energies. This state however is not a stationary state, but a superposition of stationary states, which has a probability distribution that changes with time.

- The number of nodes in the wave function increases with the energy. This is a general property of second order differential equations in one dimension having a local potential. This statement is a consequence of the **nodal-line theorem** or **Knotensatz** by Richard Courant[21].(see also Appendix F on p. 279)

4.5.1 Time-dependent solutions for the particle in the box

Note that the solutions for the stationary Schrödinger equation, provide us with the energy levels, but not with the dynamics of a general wave function. A general wave function has the form

$$\psi(x,t)=\sum_{n=1}^{\infty}\phi_n(x)\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n t}a_n$$

where the coefficients a_n are determined by the initial conditions of the problem.

Let us consider several situations:

- If only one eigenstate contributes, so that

  $$\psi(x,t)=\phi_n(x)\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n t}$$

  we observe that the probability density

  $$P(x,t)=\psi_n^*(x,t)\psi_n(x,t)=\phi_n^*(x)\phi_n(x)$$

  does not change with time. Note that this is also true for excited states. This is the reason that one calls these solutions stationary. The wave function itself, however, oscillates. The higher the energy the faster the oscillation.

- Consider now the superposition of the first and the second eigenstate

  $$\psi(x,t)=\phi_1(x)\mathrm{e}^{-\frac{i}{\hbar}\epsilon_1 t}a_1+\phi_2(x)\mathrm{e}^{-\frac{i}{\hbar}\epsilon_2 t}a_2$$

The probability density is no more constant in time

$$P(x,t) = \psi^*(x,t)\psi(x,t) = \Big(\phi_1^*(x)\psi_1(x)a_1^*a_1 + \phi_2^*(x)\psi_2(x)a_2^*a_2\Big)$$
$$+2\mathrm{Re}\Big(\phi_1^*(x)\phi_2(x)a_1^*a_2\,\mathrm{e}^{-\frac{i}{\hbar}(\epsilon_1-\epsilon_2 t}\Big)$$

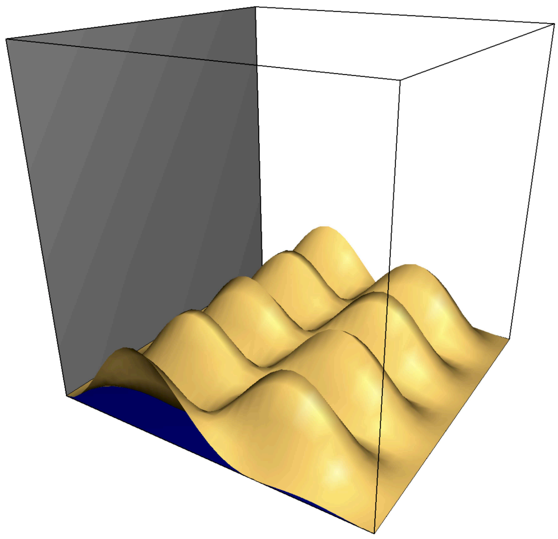


Fig. 4.6: Probability density P(x,t) for a superposition of ground and first excited state of the particle in the box as function of time. The time axis points from the front to the back.

The term that does not vary in time is mirror-symmetric, while the oscillating term is antisymmetric. This indicates that the particle oscillates back an forth with circular frequency given by $\hbar\omega = \epsilon_2 - \epsilon_1$. If the particle is charged, the particle acts like an antenna and it sends out electromagnetic radiation, i.e. light. The circular frequency of the radiation is ω, which indicates that light is emitted only with certain spectral lines corresponding to the energy differences if the discrete energy eigenlevels. Thus the system looses energy until it ends up in a pure eigenstate, which in this case is the ground state.[13] This is a first indication that atoms and molecules end up after a certain while in pure eigenstates of the Hamilton operator.

- How does the quantum mechanical result relate to the classical result? Let us create a wave packet and follow its motion.

  We start from the general Ansatz for the wave function as superposition of the solutions of the time-independent Schrödinger equation. The time dependence is included by attaching the

[13] This is not obtained from the Schrödinger equation in the present form, because the electromagnetic fields do not enter the Hamiltonian.

phase $e^{-i\omega_n t} = e^{-\frac{i}{\hbar}\epsilon_n t}$ to each eigenstate $\phi_n(x)$.

$$
\begin{aligned}
\psi(x,t) &= \sum_{n=1}^{\infty} \underbrace{\sqrt{\frac{2}{L}}\sin(k_n x)}_{\phi_n(x)} e^{-i\omega(k_n)t} a_n \\
&= \sum_{n=1}^{\infty} \sqrt{\frac{2}{L}}\frac{1}{2i}\left(e^{i(k_n x-\omega(k_n)t)} - e^{-i(k_n x+\omega(k_n)t)}\right) a_n \\
&= \underbrace{\sum_{n=1}^{\infty} e^{i(k_n x-\omega(k_n)t)}\frac{1}{i\sqrt{2L}}a_n}_{\psi_1(x,t)} - \underbrace{\sum_{n=1}^{\infty} e^{-i(k_n x+\omega(k_n)t)}\frac{1}{i\sqrt{2L}}a_n}_{\psi_2(x,t)}
\end{aligned}
$$

The solution seems to consist of two wave packets, $\psi_1(x,t)$ and $\psi_2(x,t)$, running in opposite directions. There is even more to it: we can show that each of the two solution actually is an array of partner solutions separated by a distance of $2L$. These presence of those **periodic images** is due to the equispaced discrete grid of wave vectors.

$$
\begin{aligned}
\psi_1(x,t) &= \sum_{n=1}^{\infty} e^{i(k_n x-\overbrace{2k_n L}^{-2\pi n}-\omega(k_n)t)}\frac{1}{i\sqrt{2L}}a_n = \sum_{n=1}^{\infty} e^{i(k_n(x-2L)-\omega(k_n)t)}\frac{1}{i\sqrt{2L}}a_n \\
&= \psi_1(x-2L,t)
\end{aligned}
\tag{4.21}
$$

In our problem, the periodic images are not real because at most one of them is located in the box. The other partners are only mathematical artifacts. Nevertheless they allow us to imagine the processes that occur.

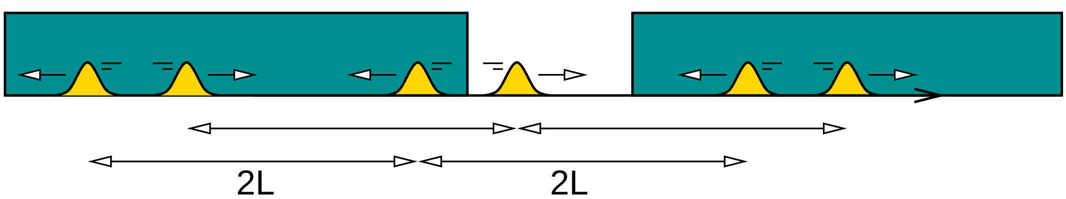
The two solutions $\psi_1(x,t)$ and $\psi_2(x,t)$ are also related by a geometric transformation: they are mirror images at the mirror plane at the right wall of the box, $x=0$.

$$
\psi_2(x,t) = \sum_{n=1}^{\infty} e^{-i(k_n x+\omega(k_n)t)}\frac{1}{i\sqrt{2L}}a_n = \sum_{n=1}^{\infty} e^{i(k_n(-x)-\omega(k_n)t)}\frac{1}{i\sqrt{2L}}a_n = \psi_1(-x,t) \tag{4.22}
$$

Furthermore, there is a mirror plane at the right wall of the box, $x=L$.

$$
\psi_2(x-L) \stackrel{\text{Eq. 4.22}}{=} \psi_1(-x+L) \stackrel{\text{Eq. 4.21}}{=} \psi_1(-x-L)
$$

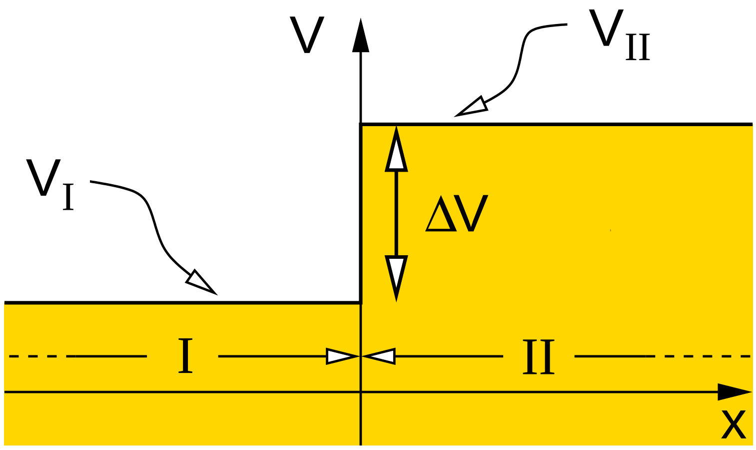
We choose the initial conditions such that the solution $\psi_1(x,t)$ is a wavepacket somewhere inside the box and with a velocity v_g. Our analysis tells us that this wave packet will move to one the walls, where it disappears "in the wall". While this wave packet disappears, one of the periodic images of $\psi_2(x,t)$ appears out "of the wall". At the walls these two solutions exactly cancel consistent with the requirement that the wave function goes to zero at the boundaries of the box. The wave packet $\psi_2(x,t)$ moves to the opposite wall, where it disappears. At the same time the next periodic image of $\psi_1(x,t)$ is ready to appear out of the wall. The result is that the wave packet bounces back and forth exactly as a classical particle would do. As the wave packet moves back an forth, it delocalizes exactly as we described it for the free particle.



4.6 Potential step and splitting of wave packets

The potential step is the basic building block for a large class of problems that can be solved fairly easily. The potential step will also make us familiar with a new quantum effect, namely the splitting of wave packets.

Imagine an interface between a metal and a semiconductor. Electrons can propagate in the semiconductor, if they have sufficient energy to enter the conduction band. Thus the electrons experience a potential step. Metal-semiconductor interface are present at electric contacts to semiconductor devices, so-called Schottky contacts. They are the key element of Schottky-diodes used as rectifiers.



The problem is defined as follows: Space is divided into two regions, region I for $x < 0$ and region II for $x > 0$. In region I we have a constant potential V_I and a solution ψ_I and in region II we have a potential V_{II} and a solution ψ_{II} so that

$$
\begin{aligned}
V(x) &= \theta(-x)V_I + \theta(x)V_{II} \\
\psi(x) &= \theta(-x)\psi_I(x) + \theta(x)\psi_{II}(x)
\end{aligned}
$$

where $\theta(x)$ is the **Heaviside function** or step function defined by

$$
\theta(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2} & \text{for } x = 0 \\ 1 & \text{for } x > 0 \end{cases} .
$$

4.6.1 Kinetic energy larger than the potential step

Let us first consider the case when the energy is sufficient to mount the step. A classical particle in such a situation will always mount the step and proceed with a smaller velocity as some of its kinetic energy has been used up to mount the step. For a quantum particle everything is different.

For a given energy, the energy eigenstates are

$$
\begin{aligned}
\psi_I(x) &= Ae^{ik_I x} + Be^{-ik_I x} \\
\psi_{II}(x) &= Ce^{ik_{II} x} + De^{-ik_{II} x}
\end{aligned}
\tag{4.23}
$$

with

$$
\hbar k_I = \sqrt{2m(E - V_I)} \qquad \text{and} \qquad \hbar k_{II} = \sqrt{2m(E - V_{II})}
\tag{4.24}
$$

For a given energy, two of the four coefficients (A, B, C, D) are determined by the condition that the wave function is continuous and differentiable at the potential step. The remaining coefficients correspond to two independent solutions of the problem, each having a variable pre-factor. To pick one solution we select $D = 0$ and $A = 1$. This corresponds to a particle beam that arrives from the left and scatters into the forward and the backward direction. No particles arrive at the step from the right side, i.e. $D = 0$.

Now we determine B and C from the requirement that the wave function is differentiable at the step.

$$\psi_I(0) = \psi_{II}(0) \overset{\text{Eq. 4.23}}{\Rightarrow} \underbrace{1}_{A} + B = C + \underbrace{0}_{D}$$

$$\partial_x\psi_I(0) = \partial_x\psi_{II}(0) \overset{\text{Eq. 4.23}}{\Rightarrow} ik_I - Bik_I = ik_{II}C$$

which results in

$$B = \frac{k_I - k_{II}}{k_I + k_{II}} \tag{4.25}$$

$$C = \frac{2k_I}{k_I + k_{II}} \tag{4.26}$$

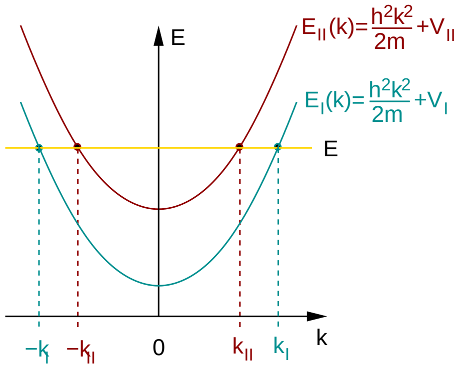


Fig. 4.7: Dispersion relations from both sides of the step. Observe that the phase velocity $v_{ph} = \frac{\omega}{k}$ in this example if higher ontop of the step than at the lower side. Nevertheless the group velocity $v_{gr} = \frac{\partial\omega}{\partial k}$ is lower ontop of the step, because some of the kinetic energy is used up by mounting the step.

Thus, all parameters are determined by the energy $E = \hbar\omega$, which in turn determines the values for $k_I(E)$ and $k_{II}(E)$ via Eq. 4.24. We obtain the time-independent solution as

$$\phi(x, E) = \theta(-x)\underbrace{\left[e^{ik_Ix} + \frac{k_I - k_{II}}{k_I + k_{II}}e^{-ik_Ix}\right]}_{\psi_I(x,\omega)} + \theta(x)\underbrace{\left[\frac{2k_I}{k_I + k_{II}}e^{ik_{II}x}\right]}_{\psi_{II}(x,\omega)} \tag{4.27}$$

The time-dependent solution is obtained by multiplication of the stationary solutions with $e^{-i\omega t}$. We obtain the time-dependent partial solutions

$$\psi_{k_I>0}(x, t) = \underbrace{\theta(-x)e^{i(k_Ix-\omega t)} + \theta(-x)\frac{k_I - k_{II}}{k_I + k_{II}}e^{i(-k_Ix-\omega t)}}_{\Psi_I} + \underbrace{\theta(x)\frac{2k_I}{k_I + k_{II}}e^{i(k_{II}x-\omega t)}}_{\Psi_{II}} \tag{4.28}$$

The solution can be characterized by the wave vector k_I of the incoming wave instead of its energy. The requirement that, in the beginning, the wave packet approaches the step from the left, translates into $k_I \geq 0$.

Already from Fig. 4.7 a lot of the relevant physics is obtained: The momentum $\hbar k$ ontop of the step is smaller than on the lower ledge. Also the velocity ontop of the step is smaller as seen from the slope of the dispersion relation at the respective wave vectors.

The striking result is however, that the particle beam is split into a transmitted beam and a reflected beam. In contrast, a classical particle either mounts the step, or, if the energy is not sufficient, it is completely reflected. A quantum particle has a certain probability to be reflected, even if the energy is sufficient to mount the step. Splitting of wave packets is one of the quantum effects.

The probability density for a beam of particles, i.e. for a plane wave, is constant in space. Thus the particle density on the right-hand side of the step is constant. On the left-hand side, the incident beam interferes with the reflected beam, leading to ripples in the density distribution. These ripples are due to standing waves. [14]

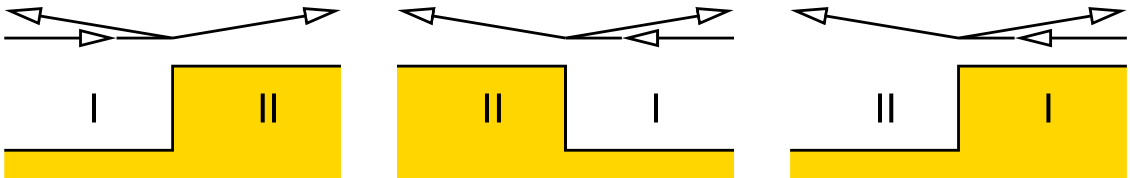
[14] A standing wave is for which the amplitude oscillates with time, but the wave crests remain in place. A standing wave is constructed from two running waves with opposite wave vectors, such as $\cos(\vec{k}\vec{r} - \omega t) + \cos(-\vec{k}\vec{r} - \omega t) = 2\cos(\vec{k}\vec{r})\cos(\omega t)$

The situation described here is rather similar to the refraction of light at the surface of an optically denser medium. Partial reflection and partial transmission at an interface is a well known observation. However some properties differ: For a particle with a mass, the momentum $\hbar k$ is smaller ontop of the barrier than on the lower ledge. For light entering a denser medium the wave vector grows larger. As a result the angle of refraction is smaller than the incident angle for light, while for particles it is larger than the incident angle.

4.6.2 Construct the other partial solution using symmetry transformations

A second-order differential equation in one dimension has always two linear independent solutions. In order to obtain the general solution, we need to construct one other independent solution of the problem.

The second partial solution can be obtained using symmetry arguments from the one just derived and presented in Eq. 4.28. Using symmetry arguments to construct solutions is an extremely powerful technique that allows one to save a lot of computations. This is why I describe this example in some detail:



The second partial solution shall describe a wave packet, which approaches the step from the right (higher) side and which hops down the step. We construct this solution by two successive transformations of the problem.

1. First we apply a mirror operation at the step. The Schrödinger equation remains independent except for the potential. The new potential $V'(x) = V(-x)$ describes a step with the upper side on the left. The partial solution given in Eq. 4.28 is transformed by interchanging x by $-x$, i.e. $x' = -x$, so that $\psi'(x,t) = \psi(-x,t)$.

$$
\begin{aligned}
\psi'(x',t) &= \psi_{k_I>0}(x,t) \\
\Rightarrow \psi'(x,t) &= \psi_{k_I>0}(-x,t) \\
\Rightarrow \psi'(x,t) &\stackrel{\text{Eq. 4.28}}{=} \underbrace{\theta(x)\mathrm{e}^{i(-k_I x-\omega t)} + \theta(x)\frac{k_I-k_{II}}{k_I+k_{II}}\mathrm{e}^{i(+k_I x-\omega t)}}_{\Psi'_{II}} + \underbrace{\theta(-x)\frac{2k_I}{k_I+k_{II}}\mathrm{e}^{i(-k_{II}x-\omega t)}}_{\Psi'_I}
\end{aligned}
$$

   We have defined $\Psi'_I = \Psi_{II}(-x,t)$ so that it describes the transformed solution on the left side of the step and $\Psi'_{II} = \Psi_I(-x,t)$ described the transformed solution on the right of the step.

   The solution obtained after this operation describes a particle approaching the step from the right-hand side, which is partially reflected and which partially hops up the step.

2. In a second operation, we restore the original potential: We interchange V_I and V_{II}, so that the elevated part of the step is again on the right hand side. We need to control that the derivation of Eq. 4.28 does not make use that $V_{II} > V_I$. The second transformation transforms $V''_I = V'_{II} = V_{II}$ and $V''_{II} = V'_I = V_I$. This implies replacing k_I and k_{II}, i.e. $k''_{II} = k_I$ and $k''_I = k_{II}$. In this way, we obtain our final result for the second partial solution.

$$
\psi''(x,t) = \underbrace{\theta(x)\mathrm{e}^{i(-k_{II}x-\omega t)} - \theta(x)\frac{k_I-k_{II}}{k_I+k_{II}}\mathrm{e}^{i(+k_{II}x-\omega t)}}_{\Psi''_{II}} + \underbrace{\theta(-x)\frac{2k_{II}}{k_I+k_{II}}\mathrm{e}^{i(-k_I x-\omega t)}}_{\Psi''_I} \quad (4.29)
$$

In summary, the second solution $\psi''(x,t)$ given in Eq. 4.29 is obtained from the first partial solution in Eq. 4.28 by the transformation

$$x' = -x \qquad \text{and} \qquad k'_{II} = k_I \qquad \text{and} \qquad k'_I = k_{II}\,.$$

This solution ψ'' describes a beam of particles approaching from the right, that is partially reflected and partially transmitted down the potential step. Surprisingly, we observe that the particle are partially reflected even at a downhill step.

Scattering of wave packets

In order to explore the physical content of these solutions, let us consider the scattering of a wave packet. We may choose one that approaches the step from the left side with coefficients $a(k)$.

$$\psi_1(x,t) = \chi(x,t)e^{i(\bar{k}_I x - \omega(\bar{k}_I)t)} = \int \frac{dk}{2\pi}\, a(k)e^{i(kx-\omega(k)t)} \tag{4.30}$$

$\psi_1(x,t)$ describes the initial wave packet before it arrives at the potential step.

If the wave packet is localized initially in region I, it can be expressed entirely by partial solutions given in Eq. 4.28. The wave packet can then be expressed by three wave packets, which we denote as Ψ_1, Ψ_2 and Ψ_3. The first one is, per construction, our initial wave packet Eq. 4.30.

$$\begin{aligned}
\psi(x,t) &\stackrel{\text{Eq. 4.28}}{=} \int \frac{dk_I}{2\pi} a(k_I)\Big[\theta(-x)e^{i(k_I x-\omega t)} + \theta(-x)B(k_I)e^{i(-k_I x-\omega t)} + \theta(x)C(k_I)e^{i(k_{II}x-\omega t)}\Big] \\
&= \theta(-x)\underbrace{\int \frac{dk_I}{2\pi}\, a(k_I)e^{i(k_I x-\omega t)}}_{\psi_1} + \theta(-x)\underbrace{\int \frac{dk_I}{2\pi}\, a(k_I)B(k_I)e^{i(-k_I x-\omega t)}}_{\psi_2} \\
&+ \theta(x)\underbrace{\int \frac{dk_I}{2\pi}\, a(k_I)C(k_I)e^{i(k_{II}x-\omega t)}}_{\psi_3} \\
&= \theta(-x)\psi_1(x,t) + \theta(-x)\psi_2(x,t) + \theta(x)\psi_3(x,t)
\end{aligned} \tag{4.31}$$

The coefficients $B(k_I)$ and $C(k_I)$ are those of Eq. 4.25 and Eq. 4.26. The wave packet $\psi_1(x,t)$ describes the incoming particles. The wave packed $\psi_2(x,t)$ describes the reflected wave particles and the wave packet $\psi_3(x,t)$ describes the particles transmitted through the barrier.

In order to simplify the discussion, let us choose a very smooth envelope function for the wave packet so that the wave function coefficients $a(k)$ are localized near a given wave vector $\bar{k}$. Then I can approximate $B(k)$ and $C(k)$ by the value at $\bar{k}$.

$$\begin{aligned}
\psi_2(x,t) &\stackrel{\text{Eq. 4.31}}{=} \int \frac{dk_I}{2\pi} a(k_I)B(k_I)e^{i(-k_I x-\omega t)} \\
&\approx B(\bar{k}_I)\int \frac{dk_I}{2\pi} a(k_I)e^{i(-k_I x-\omega t)} \\
&= B(\bar{k}_I)\psi_1(-x,t) \stackrel{\text{Eq. 4.25}}{=} \frac{\bar{k}_I - \bar{k}_{II}}{\bar{k}_I + \bar{k}_{II}}\psi_1(-x,t)
\end{aligned} \tag{4.32}$$

Thus, the reflected wave packet is identical to the incoming wave packet, but mirrored at the step and reduced in size by the prefactor $B(\bar{k}_I)$. While the incoming wave disappears at the step due to the step function, the reflected wave appears "from under its own step function".

Let us now consider the third wave packet $\psi_3(x,t)$: Here the wave vector k_{II} occurs in the plane wave. Since both, k_I and k_{II}, are linked to the energy, one can also express k_{II} as function of k_I. We use Eq. 4.24

$$\begin{aligned}
\hbar^2 k_{II}^2 = 2m(E - V_{II}) \qquad &\wedge \qquad \hbar^2 k_I^2 = 2m(E - V_I) \\
\Rightarrow\ \hbar^2 k_{II}^2 - \hbar^2 k_I^2 &= -2m(V_{II} - V_I) = -2m\,\Delta V \\
\Rightarrow\ k_{II}(k_I) &= \sqrt{k_I^2 - 2m\,\Delta V/\hbar^2}
\end{aligned} \tag{4.33}$$

Because only values of k_I close to $\bar{k}_I$ contribute appreciably to the integral, we replace this function by its Taylor expansion about $\bar{k}_I$:

$$
\begin{aligned}
k_{II}(k_I) &= \underbrace{k_{II}(\bar{k}_I)}_{\bar{k}_{II}} + \underbrace{\left.\frac{\partial k_{II}}{\partial k_I}\right|_{\bar{k}_I}}_{=\bar{k}_I/\bar{k}_{II}} (k_I - \bar{k}_I) + O(k_I - \bar{k}_I)^2 \\
&\stackrel{\text{Eq. 4.35}}{=} \bar{k}_{II} + \frac{\bar{k}_I}{\bar{k}_{II}}(k_I - \bar{k}_I) + O(k_I - \bar{k}_I)^2
\end{aligned}
\tag{4.34}
$$

We inserted here the derivative

$$
\left.\frac{\partial k_{II}}{\partial k_I}\right|_{\bar{k}_I} \stackrel{\text{Eq. 4.33}}{=} \frac{1}{2}\frac{2\bar{k}_I}{\sqrt{\bar{k}_I^2 - 2m\Delta V/\hbar^2}} \stackrel{\text{Eq. 4.33}}{=} \frac{\bar{k}_I}{k_{II}(\bar{k}_I)} \stackrel{\bar{k}_{II} \stackrel{\text{def}}{=} k_{II}(\bar{k}_I}{=} \frac{\bar{k}_I}{\bar{k}_{II}}
\tag{4.35}
$$

We use this expression in the following when we determine the relation between ψ_3 and ψ_1.

$$
\begin{aligned}
\psi_3(x,t) &\stackrel{\text{Eq. 4.31}}{=} \int \frac{dk_I}{2\pi} a(k_I) C(k_I) \mathrm{e}^{i(k_{II}(k_I)x - \omega t)} \\
&\approx C(\bar{k}_I) \int \frac{dk_I}{2\pi} a(k_I) \mathrm{e}^{i\left(\overbrace{\left[\bar{k}_{II} + \frac{\bar{k}_I}{\bar{k}_{II}}(k_I - \bar{k}_I)\right]}^{\approx k_{II}(k_I)} x - \omega t\right)} \\
&= C(\bar{k}_I) \mathrm{e}^{i(\bar{k}_{II} - \frac{\bar{k}_I}{\bar{k}_{II}}\bar{k}_I)x} \underbrace{\int \frac{dk_I}{2\pi} a(k_I) \mathrm{e}^{i(\frac{\bar{k}_I}{\bar{k}_{II}} k_I x - \omega t)}}_{=\psi_1(\frac{\partial k_{II}}{\partial k_I}x, t)}
\end{aligned}
\tag{4.36}
$$

Now we represent the incoming wave packet by a plane wave part and an envelope function

$$
\psi_1(x,t) \stackrel{\text{Eq. 4.30}}{=} \chi(x,t) \mathrm{e}^{i(\bar{k}_I x - \omega t)}
\tag{4.37}
$$

$$
\begin{aligned}
\psi_3(x,t) &\stackrel{\text{Eq. 4.36}}{=} C(\bar{k}_I) \mathrm{e}^{i(\bar{k}_{II} - \frac{\bar{k}_I}{\bar{k}_{II}}\bar{k}_I)x} \psi_1(\frac{\bar{k}_I}{\bar{k}_{II}}x, t) \\
&\stackrel{\text{Eq. 4.37}}{=} C(\bar{k}_I) \mathrm{e}^{i(\bar{k}_{II} - \frac{\bar{k}_I}{\bar{k}_{II}}\bar{k}_I)x} \chi(\frac{\bar{k}_I}{\bar{k}_{II}}x, t) \mathrm{e}^{i(\bar{k}_I \frac{\bar{k}_I}{\bar{k}_{II}}x - \omega t)} \\
&= C(\bar{k}_I) \chi(\frac{\bar{k}_I}{\bar{k}_{II}}x, t) \mathrm{e}^{i(\bar{k}_{II}x - \omega t)} \\
&\stackrel{\text{Eq. 4.26}}{=} \frac{2\bar{k}_I}{\bar{k}_I + \bar{k}_{II}} \chi(\frac{\bar{k}_I}{\bar{k}_{II}}x, t) \mathrm{e}^{i\bar{k}_{II}x - \omega t}
\end{aligned}
\tag{4.38}
$$

Thus the transmitted wave packet ψ_3 can again be related to the incoming wave packet ψ_1 by a few transformations:

- The leading wave $\bar{k}_I$ vector is replaced by a smaller one, namely $\bar{k}_{II}$.
- The envelope function $\chi(x,t)$ is compressed by the factor $\bar{k}_{II}/\bar{k}_I$
- The velocity of the wave packet is reduced by the same factor $\bar{k}_{II}/\bar{k}_I$.
- The amplitude is reduced by the factor $C \stackrel{\text{Eq. 4.26}}{=} \frac{2\bar{k}_I}{\bar{k}_I + \bar{k}_{II}}$

Thus we can obtain the wave packets Ψ_2 and Ψ_3 from Ψ_1 by simple geometric operations and a rescaling of the amplitude and main wave vector. Ψ_2 is a simple mirror image of Ψ_1 and approaches the step from the right side. Ψ_3 is located on the same side as Ψ_1 but it is stretched. While Ψ_3 has about the same envelope function as Ψ_1, the envelope function modulates a different plane wave.

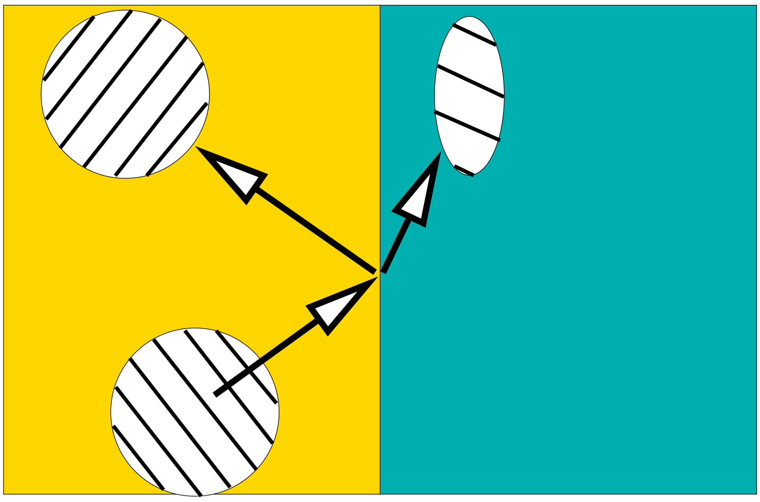


Fig. 4.8: schematic representation of a 2-dimensional wave packet scattering at a potential step. The lines represent the wave crests. The white areas represent the shape of the wave packet. The arrows indicate the velocity, respectively the wave vector.

Initially the wave function Ψ is equal to Ψ_1 and it is localized in the left half plane. Ψ_2 and Ψ_3 do not appear yet because they are localized in that half plane, where the corresponding step function vanishes. The wave packet approaches the step from the left-hand side. When it hits the step, Ψ_1 starts to disappear under its step function. At the same time the wave packets Ψ_2 and Ψ_3 appear from under their step functions, so that Ψ_1 is gradually replaced by Ψ_2 and Ψ_3. The wave packet Ψ_2 describes the particles reflected to the left, while Ψ_3 describes the particles transmitted to the right.

Reflection and transmission coefficient

The weight of the reflected wave packet is called the **reflection coefficient**

$$R \stackrel{\text{def}}{=} \int dx\ \psi_2^*(x)\psi_2(x) \stackrel{\text{Eq. 4.32}}{=} \frac{|\bar{k}_I - \bar{k}_{II}|^2}{|\bar{k}_I + \bar{k}_{II}|^2}$$

It is the probability that a particle is reflected at the potential step.

The probability that the wave packet transmitted is called **transmission coefficient**.

$$T \stackrel{\text{def}}{=} \int dx\ \psi_3^*(x)\psi_3(x) = |C(\bar{k}_I)|^2 \int dx\ \psi_1^*(\frac{\bar{k}_I}{\bar{k}_{II}}x)\psi_1(\frac{\bar{k}_I}{\bar{k}_{II}}x) = |C(\bar{k}_I)|^2 \frac{\bar{k}_{II}}{\bar{k}_I} \int dx\ \psi_1^*(x)\psi_1(x)$$
$$= \frac{|\bar{k}_{II}|}{|\bar{k}_I|} \underbrace{\frac{|2\bar{k}_I|^2}{|\bar{k}_I + \bar{k}_{II}|^2}}_{|C(\bar{k}_I)|^2} = \frac{4|\bar{k}_I||\bar{k}_{II}|}{|\bar{k}_I + \bar{k}_{II}|^2}$$

It is easily seen that $R + T = 1$, that is the particle is either transmitted or reflected. R/T is small when the kinetic energy of the incoming wave packet is much larger than the size of the step. Thus we recover the classical result in this limit. When the energy is barely sufficient to mount the step, that is when k_{II} approaches zero, the reflection coefficient becomes unity.

Thus the wave packet splits at the step in two pieces. Such a behavior is impossible for a classical particle, which always maintains its integrity. If its energy is sufficient it will always climb up the step, and does not reflect. For a wave it is not surprising it can be split. A beam of light, that is electromagnetic waves, can be partially reflected and partially transmitted through a window of glass.

The reflection and transmission coefficients as function of the energy is shown in Fig. 4.9.

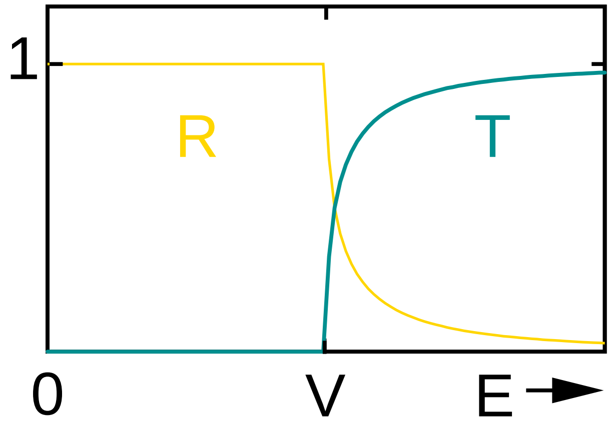


Fig. 4.9: Transmission (green) and reflection coefficient (golden) for a step of height V as function of energy. The transmission coefficient vanishes when the energy is lower than the step height and it approaches unity for high energies.

What does it mean when the wave packet splits? The particle is not really divided into two distinct pieces. The wavefunction provides us only with the probability to find the particle somewhere in space. Hence a splitting wave packet tells us that the particle has a finite probability to pass on or to be reflected.

There is one open question: Why do we not observe this beam splitting for macroscopic particles? The answer lies in the abruptness of the potential step. A macroscopic potential step is not abrupt. If the wave packet is small compared to the region over which the potential changes from one value to another, the transmission coefficient approaches one and the wave packet behaves like a classical particle. [15]

Interestingly also a wave packet hopping down the step is reflected with finite probability.

It is possible to evaluate the transmission and reflection coefficients without using wave packets. What we need are the amplitudes and wave vectors of the transmitted and reflected waves, that is Eqs. 4.24,4.25 and 4.26. We can evaluate the particle current as density times velocity, that is

$$j(x) = \Psi^*(x)\Psi(x)\underbrace{\frac{\partial\omega}{\partial k}}_{v_g} = \Psi^*(x)\Psi(x)\frac{\hbar k}{m}$$

Thus we can define the reflection coefficient as the ratio of reflected versus incoming current and the transmission coefficient as the ratio of transmitted versus incoming current.

$$T = \frac{j_3}{j_1} = \frac{|C|^2 k_{II}}{k_I} = \frac{4k_I k_{II}}{(k_I + k_{II})^2}$$
$$R = \frac{j_2}{j_1} = |B|^2 = \frac{(k_I - k_{II})^2}{(k_I + k_{II})^2}$$

Note that both, k_I and k_{II} are functions of the energy, and therefore transmission and reflection coefficient are functions of the energy as well.

[15] This can be seen as follows: we divide the potential step into many smaller steps, so that the total height is preserved. Each plateau of the potential must be much larger than the size of the wave packet. The total transmission coefficient is then the product of the transmission coefficients of the smaller steps. If the step is divided into many small steps, the total transmission coefficient tends to one. (Hint: Use steps $k_i = k_I e^{\alpha i/n}$ with $e^{\alpha} = k_{II}/k_i$)

4.6.3 Kinetic energy smaller than the potential step

If the energy is smaller than the potential step, a classical particle is always reflected at the step. We will see that a quantum particle behaves similarly. However, it can enter the classically forbidden region to a certain extent.

When the step is larger than the kinetic energy, $k_{II} = i\lambda$ is imaginary so that the plane wave is replaced by a exponential function. Thus, the solution in region II can be written as

$$\psi_{II}(x) = C\mathrm{e}^{-\lambda x} + D\mathrm{e}^{\lambda x}$$

We can immediately exclude a solution with $D \neq 0$, because the wave function is not normalizable, since it grows to infinity for large x. Thus we conclude that the wave function decays exponentially in region II.

On the left-hand side we can match a cosine function[16], that is $\psi_I(x) = A\cos(kx - \delta)$. From the matching condition for value and derivative, we obtain two equations

$$A\cos(-\delta) = C \tag{4.39}$$
$$-kA\sin(-\delta) = -C\lambda \tag{4.40}$$

for the coefficients A, B and δ. We can choose one parameter freely, so we set $A = 1$.

$$\delta = -\mathrm{atan}(\frac{\lambda}{k})$$

It is not completely straightforward to work out C from this starting point. We first determine $\cos(\delta)$ from the expression Eq. 4.41 for δ. Then we obtain C from Eq. 4.39.

$$\begin{aligned}
& & \delta &= -\mathrm{atan}(\frac{\lambda}{k}) \qquad \Rightarrow \qquad \tan(-\delta) = \frac{\lambda}{k} \\
\Rightarrow & & \frac{\lambda^2}{k^2} = \tan^2(-\delta) &= \frac{\sin^2(-\delta)}{\cos^2(-\delta)} = \frac{1-\cos^2(-\delta)}{\cos^2(-\delta)} \\
\Rightarrow & & 1-\cos^2(-\delta) &= \frac{\lambda^2}{k^2}\cos^2(-\delta) \\
& \Rightarrow & 1 &= \left(1 + \frac{\lambda^2}{k^2}\right)\cos^2(-\delta) \\
& \Rightarrow & C &\overset{\text{Eq. 4.39}}{=} A\cos(-\delta) = \sqrt{\frac{1}{1+\frac{\lambda^2}{k^2}}}
\end{aligned}$$

Thus we obtain

$$C = \sqrt{\frac{k^2}{k^2+\lambda^2}}$$

Therefore, we find

$$\psi_I(x) = \cos[kx + \mathrm{atan}(\frac{\lambda}{k})]$$
$$\psi_{II}(x) = \sqrt{\frac{k^2}{k^2+\lambda^2}}\mathrm{e}^{-\lambda x}$$

Important is that the wave function has an exponential tail that enters the classically forbidden region, even though only for a small distance. This implies that the particle has a finite probability

[16]We could also have chosen a superposition of an incoming and a reflected plane wave with complex coefficients. However, we can choose on the right hand side a real dependence, i.e. $\mathrm{Im}[C] = 0$. Then also $\psi_I(x)$ is real. The real part of the two partial solutions has the general form of a cosinus function with a phase shift.

to enter the classically forbidden region. For a classical particle this is impossible. The same effect will lead us later to the effect of tunneling: A quantum particle can even penetrate a hard wall if it is sufficiently thin.

The exponential damping of the wave function in region II corresponds to total reflection of light. If we try to look through the water surface into the pond, there is a viewing certain angle where the water surface reflects completely and all that is underneath the surface is hidden from us. If we stand up and peek into the water at a steeper angle, we suddenly observe that the water becomes more transparent.

4.6.4 Logarithmic derivative

Before we continue, we will learn a useful way to solve matching problems. The general problem is one of two regions in a one-dimensional space. We assume that we have only one solution ψ_3 in region II, and the superposition of two solutions ψ_1 and ψ_2 in region I. Without limitation of generality we choose region I to be the region left of $x = 0$ and region II to be the region right of $x = 0$.

In the case of the potential step we have selected one outgoing wave $\psi_3 = e^{ikx}$ in region II, but no incoming wave. In region I we have a superposition of ψ_1 and ψ_2. The Ansatz for the wave function is

$$\Psi(x) = \theta(-x)\Big[A\Psi_1(x) + B\Psi_2(x)\Big] + \theta(x)\Psi_3(x) \tag{4.41}$$

Without loss of generality, I have chosen $x = 0$ as the boundary between the regions I ($x < 0$) and II ($x > 0$). $\theta(x)$ is the Heaviside function with $\theta(x < 0) = 0$ and $\theta(x > 0) = 1$. Thus $\theta(x)$ equals one for region II and vanishes in region I, while $\theta(-x)$ equals one for region I and vanishes for region II.

The matching problem has the following form, where we abbreviate $\psi_i \stackrel{\text{def}}{=} \psi_i(0)$ and $\partial_x\psi_i \stackrel{\text{def}}{=} \partial_x|_0\psi_i$.

$$\begin{pmatrix} \psi_1 & \psi_2 \\ \partial_x\psi_1 & \partial_x\psi_2 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \psi_3 \\ \partial_x\psi_3 \end{pmatrix}$$

where values and derivatives are taken at the boundary $x = 0$ between the two regions I and II. After multiplication with the inverse of the matrix on the left hand side, we obtain the coefficients A and B

$$\begin{aligned} \begin{pmatrix} A \\ B \end{pmatrix} &= \frac{1}{\psi_1\partial_x\psi_2 - \psi_2\partial_x\psi_1} \begin{pmatrix} \partial_x\psi_2 & -\psi_2 \\ -\partial_x\psi_1 & \psi_1 \end{pmatrix} \begin{pmatrix} \psi_3 \\ \partial_x\psi_3 \end{pmatrix} \\ &= \frac{1}{\psi_1\psi_2(D_2 - D_1)} \begin{pmatrix} \psi_2\psi_3(D_2 - D_3) \\ \psi_1\psi_3(D_3 - D_1) \end{pmatrix} \end{aligned} \tag{4.42}$$

where we used the shorthand notation $D_i = D_{\psi_i}(0)$ for the logarithmic derivative defined below.

Definition 4.1 *LOGARITHMIC DERIVATIVE*

$$D_\psi(x_0) = \frac{\partial_x|_{x_0}\,\psi(x)}{\psi(x_0)} = \partial_x|_{x_0} \ln[\psi(x)] \tag{4.43}$$

The logarithmic derivative is independent of a scale factor, that is of the normalization.

Thus we find the solution[17] for the solution that is valid across the boundary.

$$\psi \stackrel{Eqs.\ 4.41,4.42}{=} \psi_3(0)\theta(-x)\Big[\frac{\psi_1(x)}{\psi_1(0)}\frac{D_2 - D_3}{D_2 - D_1} + \frac{\psi_2(x)}{\psi_2(0)}\frac{D_1 - D_3}{D_1 - D_2}\Big] + \theta(x)\psi_3(x) \tag{4.44}$$

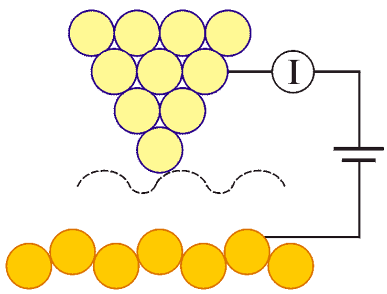
[17] It is not required to memorize this equation itself, but only its derivation.

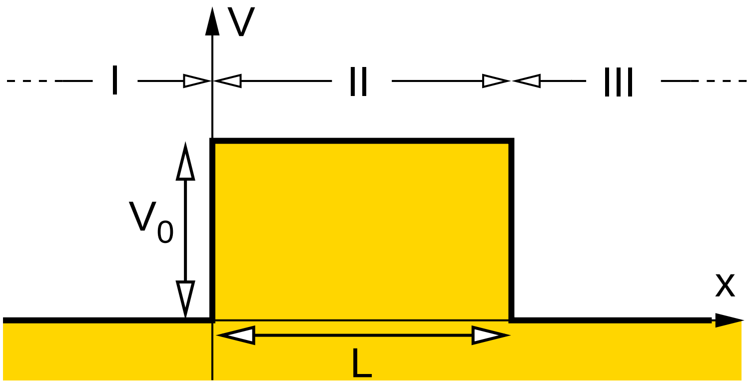
Using this formula we can easily do the matching for a system with many potential steps. One simply chooses one partial solution on one end of the steps, which plays the role of $\psi_3(x)$. Then one matches the solution of the next segment to $\psi_3(x)$. The new solution in this segment now plays the role of $\psi_3(x)$ to do the matching across the next boundary. In this ways one can proceed step by step from one end to the other and thus obtains one partial solution for the entire problem.

4.7 Barrier and tunnel effect

Let us consider a particle that approaches a barrier of finite width. If a classical particle approaches such a situation, it will overcome the barrier, if its kinetic energy is sufficiently large, and if its energy is lower than the barrier, it will be reflected. There is no grey zone. However, we have seen for the potential step, that the particle can penetrate a little into the step, even if its energy is insufficient to mount the step. This implies that if the barrier is sufficiently thin, there is a certain probability that the particle appears on the other side of the step, even if that is classically forbidden. This is the essence of the **tunnel effect**, which we will investigate here in some detail.

A famous application of the tunnel effect is the **scanning tunneling microscopy**[18] , which is one of the very few experimental techniques that can make individual atoms visible. In the scanning tunneling microscope a very fine needle is brought into contact with a surface. If a voltage is applied between the needle and the surface a current can flow. The distance between the tip of the needle and the surface is of the order of a nanometer. Due to the gap between surface and tip electrons cannot flow classically. However they can tunnel from surface to the tip. The resulting current is extremely sensitive to the distance, which is exploited to measure the topography of the surface and to control the motion of the tip.





In the following, we will determine the quantum mechanical behavior of particles impinging on a barrier. The potential has the following form

$$V(x) = \left\{ \begin{array}{c} V_0 \\ 0 \end{array} \right\} \text{for} \left\{ \begin{array}{c} 0 < x < L \\ x < 0; x > L \end{array} \right\}$$

We call the region to the left of the barrier region I, the barrier itself region II and what is to the right of the barrier is region III.

[18]Heinrich Rohrer and Gerd Binnig received, together with Ernst Ruska, the Nobel prize in Physics 1986 for the invention of the Scanning tunneling microscope.

The wave number of the particle in region I and III is $k_I \stackrel{V_I=0}{=} \frac{1}{\hbar}\sqrt{2mE}$. Two different cases need to be distinguished in the middle region II:

- Passing the barrier is classically allowed. That is the case of the kinetic energy in region I is larger than the barrier V_0. The wave function is oscillatory in the barrier region with a wave vector $k_{II} = \frac{1}{\hbar}\sqrt{2m(E-V_{II})} = \frac{1}{\hbar}\sqrt{2m(E-V_0)}$
- Passing the barrier is classically forbidden. In this case, the energy is below V_{II} and the kinetic energy in the barrier region is negative! A classical particle would be reflected by the barrier. Instead of an oscillating solution we obtain a wave function that decays (or increases) exponentially with a decay constant $\lambda = \frac{1}{\hbar}\sqrt{2m(V_0-E)}$.

We begin with the case where the kinetic energy of the particle is sufficient to mount the barrier. We only sketch the outline of the derivation and leave the detailed derivation to appendix D.1 on p. 271. We form an Ansatz of the wave function

$$\psi(x) = \theta(-x)\psi_I(x) + \theta(x)\theta(L-x)\psi_{II}(x) + \theta(x-L)\psi_{III}(x)$$

with the partial solutions for the three regions:

$$\begin{aligned}
\psi_I(x) &= A_I e^{ik_I x} + B_I e^{-ik_I x} \\
\psi_{II}(x) &= A_{II} e^{ik_{II} x} + B_{II} e^{-ik_{II} x} \\
\psi_{III}(x) &= A_{III} e^{ik_I x} + B_{III} e^{-ik_I x}
\end{aligned}$$

The Ansatz has six unknowns.

- Four of the six unknowns are determined by the conditions that the wave function must be, at each boundary, continuous with value and derivative. We call a wave function that is continuous with value and derivative, **differentiable**.
- The remaining two variables can be determined by fixing the initial condition of a physical problem: If we are interested in a particle beam impinging from the left onto the barrier, we would fix $A_I = 1$ and $B_{III} = 0$. It is sufficient to determine two independent sets of initial conditions. Every other solution can be formed by superposition of these two partial solutions.

Of interest is how many of the impinging electrons are transmitted. Thus we evaluate the transmission coefficient

$$T(E) \stackrel{\text{def}}{=} \left|\frac{A_{III}}{A_I}\right|^2$$

which we evaluated in the appendix in Eq. C.19 on p. 268. as

$$T(E) \stackrel{E \geq V}{=} \left[1 + \frac{V_0^2}{4E(E-V_0)}\sin^2(\underbrace{\frac{L}{\hbar}\sqrt{2m(E-V_0)}}_{k_{II}L})\right]^{-1} \tag{4.45}$$

The result is shown in Fig. 4.11.

The transmission coefficient has certain values where it becomes equal unity. Those states, respectively their energies, are called **resonances**. The resonance can be understood as follows: The wave packet on top of the barrier is partially reflected at the border of the barrier. If it is reflected twice it can positively interfere with itself. The interpretation would be that the particle is caught into a state with finite lifetime, where it remains for a while before it is transmitted again. Because it pauses on top of the barrier the probability density there is enhanced. In the resonance the particles are not reflected at all by the barrier, but every particle is transmitted.

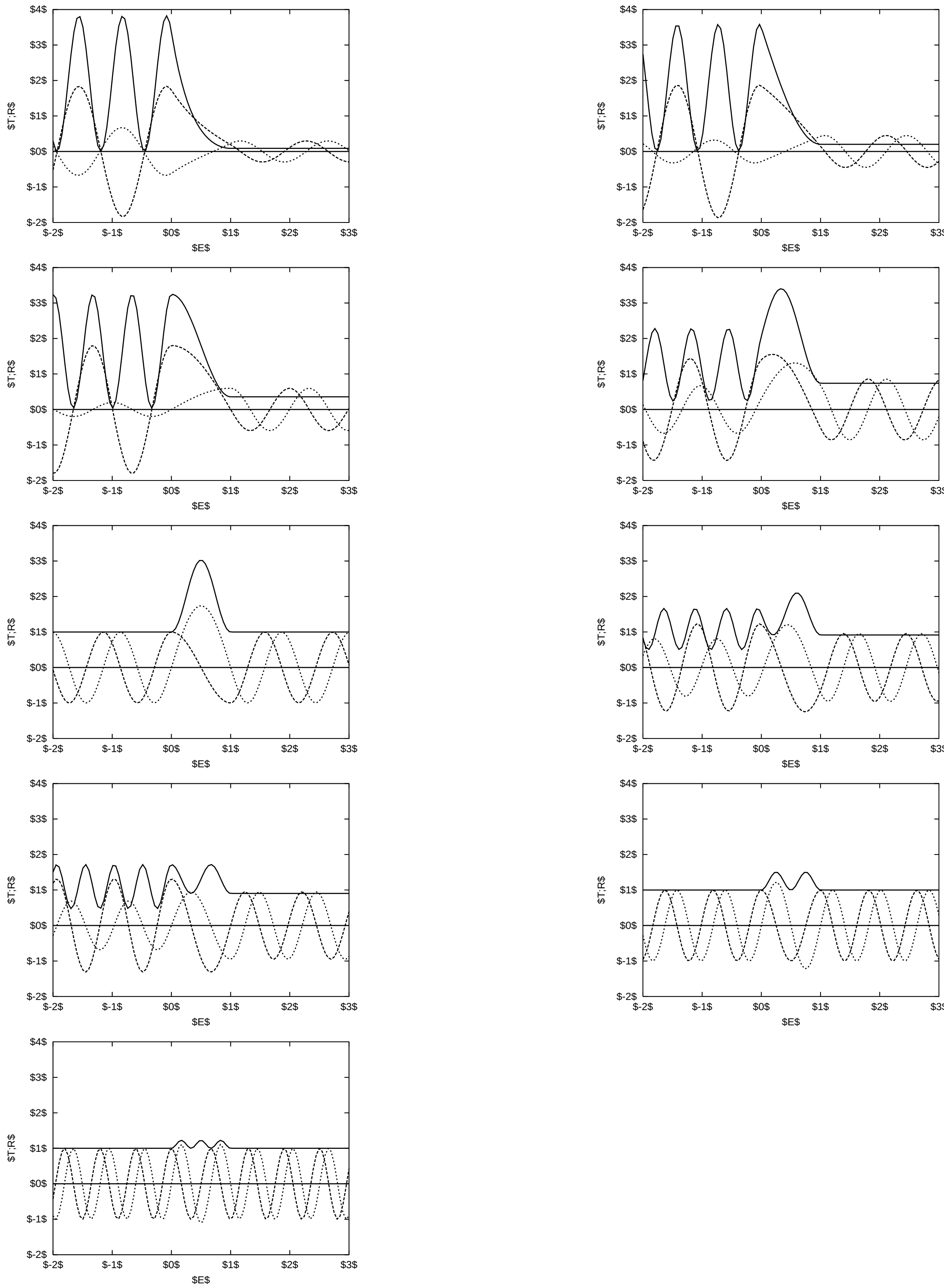


Fig. 4.10: Wave function and density of a wave scattering at a barrier of width one and height 10. The dashed line is the real part. The dotted line is the imaginary part and the full line is the probability density. The energies of the impinging particles chosen are 1 below the barrier height, and above the barrier height: 0.25, 0.5, 0.75, 1, 1.25, 1.5, 2, 3. in units of $(\hbar\pi)^2/(2mL^2)$. **Editor: more explanation!**

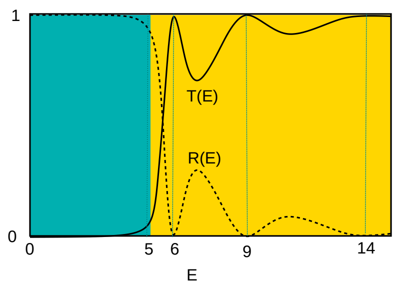


Fig. 4.11: Transmission coefficient of a barrier of height 5 and width $\pi/\sqrt{2}$. The transmission coefficient becomes nonzero, even if the energy is not sufficient to surmount the barrier classically. If the kinetic energy is larger than the barrier height, the transmission coefficient shows resonances at those energies above the barrier, where the infinite well has its eigenvalues, i.e. $E = V + \frac{(\hbar\pi)^2}{2mL^2}n^2$. The origin for these resonances lies in the fact that the particle is reflected partly even if it jumps down a potential step.

In between the resonance energies the transmission coefficient is strongly suppressed, indicating that a large fraction of the particle beam is reflected.

Let us now consider the second case, where the kinetic energy of the incoming particle is smaller than the barrier height. We can obtain the result for the transmission coefficient, by extending it into the complex plane.

In this case, we have to express

$$\sin(ix) = \frac{e^{i(ix)} - e^{-i(ix)}}{2i} = \frac{e^{-x} - e^{x}}{2i} = i\frac{e^{x} - e^{-x}}{2} = i\sinh(x)$$

so that out of Eq. 4.45 we obtain the result for $E < V$

$$T(E) = \left[1 + \frac{V_0^2}{4E(V_0 - E)}\sinh^2\left(\frac{1}{\hbar}\sqrt{2m(V_0 - E)}L\right)\right]^{-1} \tag{4.46}$$

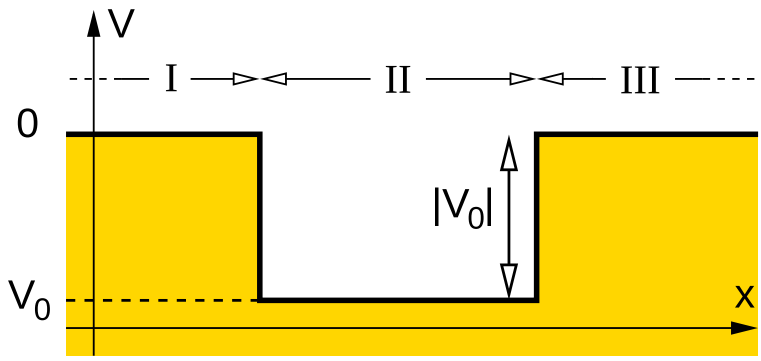
which is identical to Eq. C.20 on p. 268.

In that case, we see that the transmission coefficient Fig. 4.11 does not immediately drop to zero. Even though a classical particle with that energy would be reflected a wave packet is partly transmitted. This is the **tunnel effect**.

The tunnel effect is analogous to the observation that we can shine light through a very thin metal foil. Even though the metal should reflect light completely, some of it can pass through, if the foil is sufficiently thin.

4.8 Particle in a well and resonances

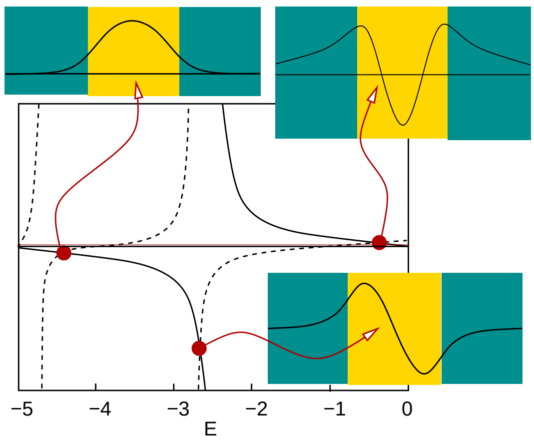
Let us consider a particle in a box with walls that allow it in principle to escape. This is a simple model for an electron bound by a nucleus. Given sufficient energy the electron may escape from the nucleus. This is the case we are considering here.



$$V(x) = \left\{ \begin{matrix} V_0 \\ 0 \end{matrix} \right\} \text{for} \left\{ \begin{matrix} 0 < x < L \\ x < 0; x > L \end{matrix} \right\}$$

If the energy of a state is smaller than zero the solutions are similar to those of the infinite box. The main difference is that the wave function does not vanish abruptly at the box boundary, but leaks out with an exponential tail, as in the finite step. As the wave function is now less constrained the energy is lowered as compared to the box with infinite walls.

Interesting are the scattering states with energy larger than zero. This is a model for an electron scattering at an atom or an impurity in a solid. We introduce three regions: Region I to the left of the well, region II is the well itself and region III lies to the right of the well.



Within each region the potential is constant so that we choose an Ansatz built up from piecewise solutions of the free particle with appropriate energy. For a given energy the wave vector in the well is $k = \pm k_0$ and the wave vector in regions I and III is $k = \pm k_1$ with k_0 and k_1 defined as

$$k_0 = \frac{1}{\hbar}\sqrt{2m(E - V_0)}$$
$$k_1 = \frac{1}{\hbar}\sqrt{2mE}$$

Note that the wave vector k can be imaginary. This happens if the energy lies below the potential energy. In that case, we expect an exponential instead of an oscillatory solution.

The detailed calculation can be found in appendix D.1. Here we only discuss the results.

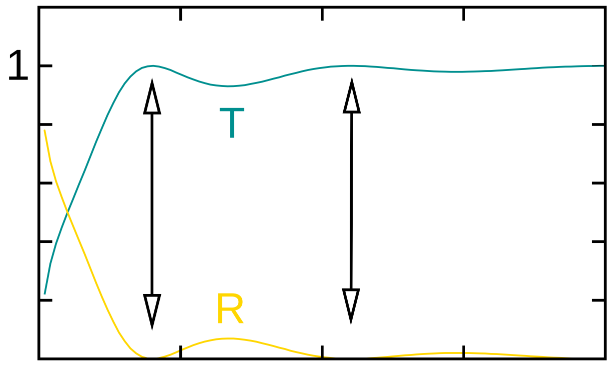


Fig. 4.12: Transmission coefficient of a potential well with depth $V_0 = -5$ and width $\pi/\sqrt{2}$. The transmission coefficient becomes unity at the resonances. The resonances, indicated by the double arrows, lie at the energies, where the box with infinite walls has its eigenvalues, i.e. $E_i = n^2 - 5$. If the kinetic energy of the particle is large compared to the well depth, the particle does not experience the well any longer and the transmission coefficient becomes unity.

We obtain the transmission coefficient $T = \left|\frac{A_{III}}{A_I}\right|^2$ as

$$T \stackrel{\text{Eq. C.19}}{=} \left[1 + \frac{V^2}{4E(E-V)}\sin^2\left(\frac{L}{\hbar}\sqrt{2m(E-V_0)}\right)\right]^{-1}$$

The transmission coefficient has maxima where $k_0 L_0 = n\pi$, n being an integer, where it is unity. This is where the infinite potential well of same width has its eigenvalues. In between it drops to a value $T = \frac{4E(E-V)}{(V-2E)^2}$, which is zero if the kinetic energy E outside the well vanishes, and which approaches one, if the kinetic energy is much larger then the well potential.

The energies, for which the transmission coefficient has its maxima, are called **resonances**. At those energies the electron is not reflected at all. The energies of the resonances can be estimated from those of the particle in a box. They are approximately at

$$E \approx V_0 + \frac{\hbar^2\pi^2}{2mL^2}n^2$$

4.9 Summary

We have learned that wave packet of a free particle becomes broader with time. This is due to the fact that probability distribution for the momenta is broad. Heisenberg's uncertainty relation indicates that the momentum distribution is broader if the position distribution is narrow. If the dispersion relation $E(p)$ is non-linear the broad momentum distribution results in a broad velocity distribution. This velocity distribution is responsible for the broadening.

An energy eigenstate for the free particle is a simple plane wave with wave vector $k = \frac{1}{\hbar}\sqrt{2mE_{\text{kin}}}$ where $E_{\text{kin}} = E - V$ is the kinetic energy. The higher the momentum of a particle, the smaller is its wave length.

For negative kinetic energy we would obtain exponentially decaying and increasing functions as eigenfunctions. They do not lead to eigenstates, because they violate the boundary conditions.

From the particle in the box we have seen that eigenstates are discrete, i.e. quantized, when the particle is enclosed in a finite region.

The lowest energy for the particle in the box is not zero, but is has a minimum energy, the zero-point energy. This observation can again be rationalized by Heisenberg's uncertainty relation, saying that a particle enclosed into a small spatial region has a large momentum uncertainty. This implies that the momentum cannot be zero, which implies that it has a finite kinetic energy. This zero-point energy becomes smaller the larger the box is. For an infinite box, that is for a free particle, the zero-point energy vanishes.

The energy eigenvalues for the particle in a box can be estimated from the quantization condition for the momenta, i.e. $k_n L = \pi n$, and the dispersion relation $E(p)$ with the de-Broglie relation $p = \hbar k$.

A wave packet arriving at a potential step with a kinetic energy, that is sufficient to classically mount the step, is split into a reflected and a transmitted wave package. If the energy is smaller than the barrier, the probability density does not disappear abruptly at the step, but the wave function exhibits an exponential tail that enters the classically forbidden region. Nevertheless, all particles are reflected. For light, the exponential tails in the classically forbidden region correspond to the so-called **evanescent waves**. If the kinetic energy is very different from the potential step, either much lower or much higher, the particle behaves classically, that is either the reflected or the transmitted wave packet vanishes.

The potential well exhibits quantized energy levels for negative energies and a continuum of scattering states for positive energies. The bound states are closely related to the states of the hard box. It differs by the fact that the wave function does not go to zero abruptly at the boundaries of the box, but that it has exponential tails reaching out into the classical forbidden region. The energy levels are slightly below those of the hard box, because they have a little more room, because they can leak out of the box a little.

For positive energies, the wave functions of the particle well are plane waves, that are modulated in the well: The wave length is shorter, due to the larger kinetic energy and the probability density may be larger. The particle is partly reflected at each boundary of the well. Hence it can, after being reflected twice, interfere constructively or destructively with the original wave. For constructive interference the probability density in the well can be much larger than outside. This is an indication of a resonance. The resonance is the relict of the bound states of the hard box. This state can be described classically by a particle that is captured in the well for a while, after which it is emitted again. The life-time of the particle in the well is related to the width of the resonance, i.e. the probability of finding the particle in the well as function of energy. The life-time is inverse proportional to the width of the resonance. This behavior is reminiscent of the behavior of a driven harmonic oscillator with friction. The width of the resonance peak in the susceptibility of a harmonic oscillator is inverse proportional to the decay time (see ΦSX:Klassische Mechanik[14]).

For the barrier, the particle will penetrate the barrier even if its energy is insufficient to classically mount the barrier. Inside the barrier, the amplitude decays exponentially. Hence there is a finite amplitude at the opposite side of the well, which leads to a transmitted wave. Hence the particle can penetrate a barrier that, classically is too high. This is the so-called tunneling effect.

If the kinetic energy is larger than the barrier, the particle is reflected on both sides of the barrier, which again leads to resonances just as in the potential well. The resonances are roughly at the location where the quantization condition for the reflected wave is met, that is at the location where the kinetic energy ontop of the barrier is related to the energy levels of the hard box.

Editorial remark: describe wave functions at the resonances, amplitude in the well, phase shifts, delay and lifetime

4.10 Recommended exercises

1. MOSFET: Exercise 16.2 on p. 243. (not ready)
2. Quantized conductance: Exercise 16.3 on p. 244. (not ready)

Chapter 5

Language of quantum mechanics

In this section we will become familiar with the mathematical basis of quantum mechanics.

Any predictive theory must have three main elements:

- **States** that describe momentary physical situations
- **Equations of motion** that describe how a physical state evolves in time. They are needed to extrapolate into the future.
- **Measurements** that allow to extract observable quantities from a state which can be compared with experiment. Measurements are needed to specify the initial state by comparison with an experimental situation, and to verify or falsify the result by comparison with experiment. Note that here measurement has the meaning of extracting a prediction on outcome of a real experiment from a state. Therefore, it is a theoretical concept.

In this and the following section we will define these ingredients for quantum theory.

Fig. 5.1: Paul Dirac, 1902-1984. British Physicist. Nobel price 1933 in Physics for the Dirac equation describing relativistic electrons and their anti-particles, the positrons.

5.1 Kets and the linear space of states

A quantum mechanical state is represented by a **ket** written as $|\psi\rangle$. A state can be expressed by a wave function in coordinate space as we have done before. There is nothing wrong to have a wave function in mind when thinking about a state. The same state can be represented either as a wave function in real space $\Psi(x)$ or, alternatively, by its Fourier transform $\Psi(k)$. Both carry the same information. We call these two ways of describing the state different **representations** of the same state. The concept of representations will be described more rigorously later. We will see that there are a number of different and equally important representations of a state, and finding a suitable representation is the key to solve many problems. The concept of a state or ket abstracts from the particular representation. A state simply stands for the physical object.

All possible states, or kets, span the **Hilbert space**[1].

The space of all kets has the same mathematical structure as the space of all **square-integrable**[2], complex functions in real space, the wave functions. It also has the same structure as the space of

[1]Strictly speaking, the set of quantum mechanical states with the operations defined below is a Hilbert space, but not the only one. Physicists often refer to the set of all quantum mechanical states as the Hilbert space. Hilbert spaces have been defined by von Neumann[22]

[2]A square integrable complex function in one dimension has a defined, finite value of the integral $\int dx\, f^*(x)f(x)$.

vectors in infinite dimensions. Thus, in the following, rather formal, description of the basic postulates of the Hilbert space, it is useful to compare and verify the relations also for square-integrable functions $\psi(x)$ and vectors $\vec{r}$.

It may be useful to compare the abstraction from wave functions to kets, by comparing it to the abstraction from n-tuples of numbers to vectors. When we begin to work with vectors, we usually use n-tuples of numbers such as (r_x, r_y, r_z) to describe a vector. Later we introduce an abstract symbol $\vec{r}$ for it. We realize that the n-tuple can represent a vector, if a coordinate system, that is three basis vectors $\vec{e}_x$, $\vec{e}_y$ and $\vec{e}_z$, is chosen. The basis vectors are considered normalized and orthogonal. That is we consider a cartesian coordinate system. We obtain the components of a vector by $r_x = \vec{e}_x\vec{r}$, $r_y = \vec{e}_y\vec{r}$ and $r_z = \vec{e}_z\vec{r}$. We can also go the other way round and compose a vector from its components as $\vec{r} = \vec{e}_x r_x + \vec{e}_y r_y + \vec{e}_z r_z$.

Now we do the analogous abstractions from wave functions to kets: We start out from a wave function $\psi(x)$. Note that a function maps the argument x to a value namely $\psi(x)$, similar as the n-tuple maps the index $i \in \{x, y, z\}$ to a number r_i. A function is a very similar object as a vector: if we make the index i continuous, we obtain a function $\psi(x)$. A more abstract description of a wave function is a ket $|\Psi\rangle$, just as the abstract notation of a vector is $\vec{r}$. In order to obtain the wave functions we need to introduce a basis. One possible basisset are the states $|x\rangle$ which correspond to δ-functions in real space. That is, the wave function related to the ket $|x_0\rangle$ is $\delta(x - x_0)$.

In order to do the next step we need to define a bra $\langle\psi|$ corresponding to a ket $|\psi\rangle$. When a ket $|\psi\rangle$ is related to a wave function $\psi(x)$ the corresponding bra $\langle\psi|$ is related to the complex conjugate wave function $\psi^*(x)$.

Bra's and ket's are related to each other like standing and lying vectors. A lying vector can be seen as an $1 \times n$ matrix and a standing vector can be seen as a $n \times 1$ matrix. A lying vector on the left multiplied with a standing vector on the right gives the scalar product of the two vectors. The scalar product is thus obtained using the usual rules of matrix multiplication. Whereas in the common vector notation we use symbols such as a dot for the scalar-product and $\otimes$ for the dyadic product, the matrix-like notation of vectors as standing and lying vectors does not require different symbols for the different products. Instead, the type of product immediately follows from the order with which the vectors are multiplied.

We proceed similarly for bra's and ket's. The scalar product between two states is defined as the product of a bra on the left and a ket on the right-hand side.

$$\langle\psi|\phi\rangle \stackrel{\text{def}}{=} \int dx\ \psi^*(x)\phi(x)$$

This expression is analogous to the definition of the scalar product of vectors as $\vec{a} \cdot \vec{b} = \sum_j a_j b_j$. The main difference is that the index j is converted into a continuous argument x and the sum is correspondingly converted into an integral.

The form of the expression is also the reason for the names **"Bra"** and **"Ket"**: A bra and a ket together make up a **"Bracket"**. Often an operator is sandwiched "bracketed" between a bra and a ket such as $\langle\psi|\hat{A}|\phi\rangle$.

Now we can obtain the "real space components" of a ket $|\psi\rangle$, namely the wave function, as

$$\langle x|\psi\rangle = \int dx'\ \delta(x - x')\psi(x') = \psi(x)$$

This expression is analogous to the way we obtain vector components as $\vec{e}_i\vec{r} = r_i$.

We can also reconstruct the ket from its wave function as

$$|\psi\rangle = \int dx\ |x\rangle\psi(x)$$

which is analogous to the vector expression $\vec{r} = \sum_i \vec{e}_i r_i$.

5.1.1 Axioms

In the following I will present the definition of the so called **Hilbert space**, which defines the basic rules for doing computations with kets. The term “Hilbert space” has been coined by von Neumann[22], who also laid down the axiomatic foundation of quantum mechanics. His classic book[23] is an excellent reference on this topic.

It is helpful to compare and verify these rules for vectors and wave functions. The analogy with the familiar expressions will help to memorize the abstract rules.

Addition of states

The addition is defined as follows:

- for each pair of states in the Hilbert space $\mathcal{H}$ the sum $|\psi_1\rangle + |\psi_2\rangle$ is a state of $\mathcal{H}$. We may also use a short hand notation $|\psi_1 + \psi_2\rangle \stackrel{\text{def}}{=} |\psi_1\rangle + |\psi_2\rangle$ for the sum.
- **Commutative Law** with respect to addition:

$$|\psi_1\rangle + |\psi_2\rangle = |\psi_2\rangle + |\psi_1\rangle \tag{5.1}$$

  We say that the addition of states is commutative
- **Associative Law** with respect to addition of states:

$$|\psi_1\rangle + \Big(|\psi_2\rangle + |\psi_3\rangle\Big) = \Big(|\psi_1\rangle + \psi_2\rangle\Big) + |\psi_3\rangle \tag{5.2}$$

  We say that the addition of states is associative.
- There is a **zero element** $|\varnothing\rangle$:

$$|\psi\rangle + |\varnothing\rangle = |\psi\rangle \tag{5.3}$$

  For teaching purposes, I have introduced the symbol $|\varnothing\rangle$ for the zero state. However, commonly one denotes the zero state simply as 0, just like the number zero. Also this notation cannot lead to misunderstandings.

  Note that the symbol $|0\rangle$ is often used denote the ground state, which is not the zero state. Because of this ambiguity I recommend to avoid using the symbol $|0\rangle$ for the zero state.
- For each state $|\psi\rangle$ there is a **negative state** $|-\psi\rangle$:

$$|\psi\rangle + |-\psi\rangle = |\varnothing\rangle \tag{5.4}$$

A set of states with the addition defined as just shown is called an **abelian group under addition**. A **abelian group** is a set with some operation that obeys the properties shown above for the addition. It is called abelian, because the commutative law holds.

Multiplication by a scalar

We define the multiplication by a scalar by

- For each state $|\psi\rangle$ in $\mathcal{H}$ and each complex number c, there is a state $|\psi\rangle c$ in $\mathcal{H}$.
- **Associative Law** with respect to multiplication by a scalar:

$$|\psi\rangle(ab) = (|\psi\rangle a)b \tag{5.5}$$

- **Distributive Law** with respect to multiplication by a scalar: The multiplication with a scalar is distributive

$$(|\psi_1\rangle + |\psi_2\rangle)c = |\psi_1\rangle c + |\psi_2\rangle c \tag{5.6}$$

We may also use a short hand notation $|\psi_1 + \psi_2\rangle = |\psi_1\rangle + |\psi_2\rangle$ for the sum and $|\psi c\rangle = |\psi\rangle c$ for the product with a scalar.

Up to now, that is after defining addition and the multiplication with a scalar, we have defined a **linear vector space**[3] over the complex numbers.

Scalar product

A linear vector space with the scalar product as defined below is a **Hilbert Space**[22, 23].

We define a scalar product $\langle\psi_1|\psi_2\rangle$ between two states $|\psi_1\rangle$ and $|\psi_2\rangle$ by

- $\langle\psi_1|\psi_2\rangle$ is a complex number.
- **Hermitian Symmetry**: Exchanging the left and right state in the scalar product yields the complex conjugate of the original scalar product.

$$\langle\psi_2|\psi_1\rangle = \langle\psi_1|\psi_2\rangle^* \tag{5.7}$$

- **Linearity**: Linearity requires distributivity Eq. 5.8, and associativity for a factor of the ket Eq. 5.9.

$$\langle\psi_1|\psi_2 + \psi_3\rangle = \langle\psi_1|\psi_2\rangle + \langle\psi_1|\psi_3\rangle \tag{5.8}$$
$$\langle\psi_1|\psi_2 c\rangle = \langle\psi_1|\psi_2\rangle c \tag{5.9}$$

  Note however, that $\langle c\psi_1|\psi_2\rangle = c^*\langle\psi_1|\psi_2\rangle$! This is a consequence of the Hermitian symmetry Eq. 5.7. The scalar product is therefore a **sesquilinear form**, which implies

- **Positive definiteness:** The scalar product of a state with itself is positive definite. It is zero exactly if the state is the zero state.

$$\langle\psi|\psi\rangle \geq 0$$

  Note that from the preceding axioms we can show that the scalar product of a state with itself is a real number, so that we can use here the greater or equal relation.

- **Completeness of the Hilbert space:** The limit of every Cauchy sequence in the Hilbert space is again a member of the Hilbert space. A **Cauchy sequence** is a sequence $|\Psi_1\rangle, |\Psi_2\rangle, \ldots$ of elements (states) with the property that

$$\langle\psi_i - \psi_j|\psi_i - \psi_j\rangle$$

  for all elements $|\psi_i\rangle$ and $|\psi_j\rangle$ in the sequence can be made smaller than any positive number ϵ, by dropping a finite number of elements from the beginning of the sequence.

- **Separability** of the Hilbert space: We will not discuss this axiom here and refer instead to the Book of von Neumann[23].

[3] A nonlinear vector space would be the space of distances in a curved space such as a sphere surface. Here for example the distributive law for the multiplication with a scalar is not valid. Take one vector from the north pole to the equator and a second one a quarter turn along the equator. If we add the two vectors we obtain a vector that points from the north pole to the equator. If we double the two vectors, the first leads to the south pole and the second moves back to the north pole. This result is clearly different than twice the sum of the two original vectors, which points from north pole to south pole.**Editor: Confirm that this is a nonlinear vector space**

5.1.2 Bra's and brackets

A scalar product is called a **bracket**. If speak about the left state in the BRAcket, we call it a **bra**. A bra is written as $\langle\psi|$. We can use the following rules for bras, which derive from the axioms mentioned above.

$$\langle\psi c| = c^*\langle\psi| \tag{5.10}$$
$$\langle\psi_1 + \psi_2| = \langle\psi_1| + \langle\psi_2| \tag{5.11}$$

5.1.3 Some vocabulary

- The **norm** of a state is

$$\sqrt{\langle\psi|\psi\rangle} \tag{5.12}$$

  the square root of the scalar product of the state with itself. Note that the scalar product of a state with itself is often called the norm, which can lead to misunderstandings.

- A state is called **normalized**, if the norm equals unity, i.e. $\langle\psi|\psi\rangle = 1$.

- Two states are called **orthogonal**, if their scalar product vanishes. That is $\langle\psi_1|\psi_2\rangle = 0$.

- A set of states $\{|\psi_i\rangle\}$ is called **orthonormal**, if $\langle\psi_i|\psi_j\rangle = \delta_{i,j}$. The states in an orthonormal set are both orthogonal to each other and they are normalized.

- A set of states $\{|\psi_i\rangle\}$ is called **linear independent**, if we can create the zero state only by the trivial superposition of those states, i.e.

$$\sum_i |\psi_i\rangle c_i = 0 \qquad \Rightarrow \qquad c_i = 0$$

- A set of states is called **linear dependent**, if they are not linear independent.

- A set of states is $\{|\psi_i\rangle\}$ called **complete** if any state $|\phi\rangle$ can be expressed as superposition of the states in the set, that is

$$|\phi\rangle = \sum_i |\psi_i\rangle c_i .$$

5.2 Operators

Operators are transformations of states like functions are transformations of numbers. An operator can be defined by defining the result of the operation for every possible state. We denote an operator by $\hat{A}$ and the result of the operation on a state $|\psi\rangle$ as $A|\psi\rangle$.

Here some analogies

- A **vector** maps (i.e. transforms) a natural number from a finite set onto a number, its corresponding vector component: $i \to y = y_i$
- a **matrix** maps a vector onto a vector: $\vec{x} \to \vec{y} = \boldsymbol{A}\vec{x}$
- a **function** maps a number onto a number: $x \to y = f(x)$
- a **functional** maps a function onto a number. $f(x) \to y = F[f(x)]$ An example of a functional is the definite integral.

- an **operator** maps a state onto a state: $|\psi\rangle \to |\phi\rangle = \hat{A}|\psi\rangle$. We will see that a state can be represented by a vector. In that case the operator will be represented as a matrix mapping a vector onto a vector. If the state is represented by a wave function, the operator will map functions onto functions. An example for an operator is the derivative, which maps a function $\psi(x)$ onto a different function $\frac{d\psi}{dx}$, namely that function which is the corresponding derivative in each point in space.

5.2.1 Axioms

Not all operators are linear, but in quantum mechanics one uses almost exclusively linear operators. **Linear** operators have the following two properties:

- $\hat{A}(|\psi_1\rangle + |\psi_2\rangle) = \hat{A}|\psi_1\rangle + \hat{A}|\psi_2\rangle$
- $\hat{A}(|\psi\rangle c) = (\hat{A}|\psi\rangle)c$

From the above definition and the axioms for states, we can derive the basic rules for the algebra of linear operators.

Addition

We begin to define the addition of operators as

$$(\hat{A} + \hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle$$

which connects the addition of operators with that of states. The addition of operators has the properties

- for any two linear operators $\hat{A}$ and $\hat{B}$ there is a linear operator $\hat{A} + \hat{B}$
- **Commutative Law** with respect to addition of linear operators

$$\hat{A} + \hat{B} = \hat{B} + \hat{A} \tag{5.13}$$

- **Associative Law** with respect to addition of linear operators:

$$(\hat{A} + \hat{B}) + \hat{C} = \hat{A} + (\hat{B} + \hat{C}) \tag{5.14}$$

- There is a **zero operator** $\hat{\emptyset}$

$$\hat{A} + \hat{\emptyset} = \hat{A} \tag{5.15}$$

- For each operator $\hat{A}$ there is a **negative operator**.

$$\hat{A} + (-\hat{A}) = \hat{\emptyset} \tag{5.16}$$

Multiplication of operators with a scalar

The rules for the multiplication with a scalar follow directly from

$$\left(\hat{A}c\right)|\psi\rangle = \left(\hat{A}|\psi\rangle\right)c \tag{5.17}$$

and the preceding axioms.

Multiplication of operators

Now we consider the multiplication of operators with themselves:

- for any two linear operators $\hat{A}$ and $\hat{B}$ there is a linear operator $\hat{A}\hat{B}$
- **Associative Law**

$$(\hat{A}\hat{B})\hat{C} = \hat{A}(\hat{B}\hat{C}) \tag{5.18}$$

- There is a **unity operator** $\hat{1}$:

$$\hat{A}\hat{1} = \hat{1}\hat{A} = \hat{A} \tag{5.19}$$

  The unit operator is the **identity**

- **Distributive Law** with respect to multiplication of two operators

$$(\hat{A} + \hat{B})\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C} \tag{5.20}$$

The rules for operators and numbers are fairly similar. However, two axioms are missing!

- Two operators do not necessarily commutate, i.e. we cannot assume $\hat{A}\hat{B} = \hat{B}\hat{A}$.
- Not every operator has an inverse.

5.2.2 Some vocabulary

- An operator $\hat{A}^\dagger$ is the **Hermitian conjugate** or **adjoint** to the operator $\hat{A}$, if , for each ket $|\psi\rangle = \hat{A}|\phi\rangle$, the corresponding bra is $\langle\psi| = \langle\phi|\hat{A}^\dagger$.

ADJOINT OPERATOR

$$|\psi\rangle = \hat{A}|\phi\rangle \qquad \Leftrightarrow \qquad \langle\psi| = \langle\phi|\hat{A}^\dagger \tag{5.21}$$

so that

$$\langle\phi|\hat{A}^\dagger|\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^* = \langle\hat{A}\phi|\psi\rangle \tag{5.22}$$

Hermitian conjugate operators are required to interchange the order of bra's and kets of matrix elements.

Let us make the connection with matrices: for a Matrix $\boldsymbol{A}$ with matrix elements $A_{i,j}$, the hermitian conjugate matrix is defined as $(A^\dagger)_{i,j} = A^*_{j,i}$ or $\boldsymbol{A}^\dagger = (\boldsymbol{A}^\top)^*$. The hermitian conjugate is obtained by transposing a matrix and taking the complex conjugate of the result.

We can easily show[4] a few important rules that can be derived from the above axioms:

$$\left(\hat{A}^\dagger\right)^\dagger = \hat{A}$$
$$(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger$$
$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$$

[4] We show that $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$

$$\langle\psi|(\hat{A}\hat{B})^\dagger|\phi\rangle \overset{\text{Eq. 5.22}}{=} \langle\phi|\hat{A}\hat{B}|\psi\rangle^* \overset{\text{Eq. 5.22}}{=} \langle\hat{A}^\dagger\phi|\hat{B}|\psi\rangle^* \overset{\text{Eq. 5.22}}{=} \langle\hat{B}^\dagger\hat{A}^\dagger\phi|\psi\rangle^* \overset{\text{Eq. 5.22}}{=} \langle\psi|\hat{B}^\dagger\hat{A}^\dagger|\phi\rangle$$

- An operator that fulfills

$$\hat{A}^\dagger = \hat{A} \tag{5.23}$$

is called **hermitian**.

Only hermitian operators can be related to observables (i.e. measurable quantities). Measurable quantities must correspond to real expectation values. Only hermitian operators have the property that all their expectation values are real.

$$\langle\psi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}^\dagger|\psi\rangle^* = \langle\psi|\hat{A}|\psi\rangle^* \tag{5.24}$$

Example: Show that the differential operator $\hat{p} = \frac{\hbar}{i}\partial_x$ is hermitian, if the scalar product is defined as $\langle\psi|\phi\rangle = \int dx\ \psi^*(x)\phi(x)$.

$$\begin{aligned}
\langle\psi|\hat{p}^\dagger|\phi\rangle = \langle\phi|\hat{p}|\psi\rangle^* &= \left(\int_{-\infty}^{\infty} dx\ \phi^*(x)\frac{\hbar}{i}\partial_x\psi(x)\right)^* \\
&= -\frac{\hbar}{i}\left(\int_{-\infty}^{\infty} dx\ \left[\partial_x(\phi^*(x)\psi(x)) - \psi(x)\partial_x\phi^*(x)\right]\right)^* \\
&= -\frac{\hbar}{i}\left(\underbrace{[\phi^*(x)\psi(x)]_{-\infty}^{\infty}}_{=0} - \int_{-\infty}^{\infty} dx\ \psi(x)\partial_x\phi^*(x)\right)^* \\
&= \frac{\hbar}{i}\int dx\ \psi^*(x)\partial_x\phi(x) = \int dx\ \psi^*(x)\frac{\hbar}{i}\partial_x\phi(x) = \langle\psi|\hat{p}|\phi\rangle \\
\Rightarrow \hat{p}^\dagger = \hat{p}
\end{aligned}$$

- The **commutator** $[\hat{A}, \hat{B}]$ of two operators is defined as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{5.25}$$

The commutator will be important because observables that are represented by operators that do not commutate, cannot be measured at the same time with arbitrary precision. This will be explained in connection with Heisenberg's uncertainty relation. Sets of commutating operators can be used as quantum numbers.

Some rules for commutators can be derived from the definition Eq. 5.25.

$$\begin{aligned}
[c\hat{A}, \hat{B}] &= c[\hat{A}, \hat{B}] \\
[\hat{A}, \hat{B}] &= -[\hat{B}, \hat{A}] \\
[\hat{A} + \hat{B}, \hat{C}] &= [\hat{A}, \hat{C}] + [\hat{B}, \hat{C}] \\
\hat{A}\hat{B} &= \hat{B}\hat{A} + [\hat{A}, \hat{B}] \\
[\hat{A}\hat{B}, \hat{C}] &= \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} \\
[\hat{A}, \hat{B}]^\dagger &= [\hat{B}^\dagger, \hat{A}^\dagger]
\end{aligned}$$

Note that the commutator is sometimes written with a minus-sign as subscript.

- Two operators with vanishing commutator are called **commutating** operators.

- A **canonical momentum** $\hat{P}$ for a generalized coordinate $\hat{Q}$ is an operator, which obeys the commutator relation

$$[\hat{P}, \hat{Q}] = \frac{\hbar}{i} \tag{5.26}$$

The generalized coordinate Q must be a hermitian operator. For example momentum and position are **canonical conjugate** of each other.

The canonical momentum p for the position x has been defined in classical mechanics by Eq. 3.5 as derivative of the Lagrange function with respect to the velocity. The physical meaning of the canonical momentum p results from the Noether theorem as conservation number related to a symmetry of the action with respect to translation in the coordinate x.

The quantum mechanical version of the canonical momentum follows from the correspondence principle, Eq. 3.28, which relates the momentum to the wave vector of a wave packet. This yields the expression $p \to \hbar k \to \frac{\hbar}{i}\partial_x$. This relation in turn determines the commutator $[p, x] = \frac{\hbar}{i}$ between the coordinate and the canonical momentum.[5]

- An **inverse** $\hat{A}^{-1}$ of an operator fulfills

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{1}$$

Not every linear operator has an inverse: For example if an operator maps two different states onto the same resulting state, we cannot find an operator, that knows from which initial state the result was obtained.

- An operator $\hat{U}$ is called **unitary** if

$$\hat{U}^\dagger\hat{U} = \hat{1} \qquad \text{or} \qquad \hat{U}^{-1} = \hat{U}^\dagger$$

If a unitary operator $\hat{U}$ has an inverse, then its inverse is its hermitian conjugate $\hat{U}^\dagger$.

While hermitian operators are important as they represent observables, unitary operators are important as they describe transformations between physical states: Such a transformation must map a normalized state onto a normalized state, which follows directly from the following theorem:

Unitary operators and only unitary operators preserve the scalar product of two states if both states are transformed by the same operation $|\phi'\rangle = \hat{U}|\phi\rangle$ and $|\psi'\rangle = \hat{U}|\psi\rangle$.

$$\langle\phi'|\psi'\rangle = \underbrace{\langle\hat{U}\phi}_{\langle\phi'|}|\hat{U}\psi\rangle = \langle\phi|\underbrace{\hat{U}^\dagger\hat{U}}_{\hat{1}}|\psi\rangle \stackrel{!}{=} \langle\phi|\psi\rangle \tag{5.27}$$

Furthermore, any transformation between states, that preserves the norm of all states is unitary.

The analogy of unitary operators are unitary matrices, which describe rotations and mirror operations of vectors. This has been discussed in *Phi*SX:Klassische Mechanik[14]. Rotations and mirror images, and translations are those transformations that preserve the distance, the norm, between any two points in space.

- We can also form **functions of operators** . The function of an operator is defined by its power series expansion (Taylor expansion): In order to work out the function of an operator, one starts from a regular function $f(x)$, which depends on numbers x. One build the Taylor expansion of the function, and then inserts the operator A for the argument.

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{1}{n!}\left(\left.\frac{d^n f}{dx^n}\right|_{x_0}\right)(x - x_0)^n \\ \Rightarrow f(\hat{A}) &= \sum_{n=0}^{\infty} \frac{1}{n!}\left(\left.\frac{d^n f}{dx^n}\right|_{x_0}\right)(\hat{A} - x_0\hat{1})^n \end{aligned} \tag{5.28}$$

[5] This argument does not prove the reverse, namely that every pair forming the commutator relation Eq. 5.26 can be mapped onto a momentum as defined in classical mechanics. This proof can be found in the article[24] by John v. Neumann.

Since additions and multiplication of operators have been defined, the function of the operator can be performed. One should remember, though, that a Taylor expansion has a finite **radius of convergence**. This implies that the result may not be defined for all operators.

The function f of an operator A can also be defined by its eigenstates

$$\hat{A}|\psi\rangle = |\psi\rangle a \qquad \Rightarrow \qquad f(\hat{A})|\psi\rangle = |\psi\rangle f(a) \tag{5.29}$$

Example for the function of a matrix

We will see in the future that operators can be represented as matrices. Let us therefore calculate a function of a matrix. We use the matrix

$$\boldsymbol{A} \widehat{=} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Let us form the exponential function of $\boldsymbol{A}$. The power series expansion of the exponential function is

$$\mathrm{e}^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$$

We insert the argument, which is now a matrix, as argument into the function and the power-series expansion.

$$\mathrm{e}^{\boldsymbol{A}t} = \sum_{n=0}^{\infty} \frac{1}{n!} (\boldsymbol{A}t)^n$$

Now we need to determine the powers of $\boldsymbol{A}$:

$$\begin{aligned}
\boldsymbol{A}^0 &= \mathbf{1} \\
\boldsymbol{A}^1 &\widehat{=} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \widehat{=} \boldsymbol{A} \\
\boldsymbol{A}^2 &\widehat{=} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \widehat{=} -\mathbf{1} \\
\boldsymbol{A}^3 &= \boldsymbol{A}^2 \boldsymbol{A} = -\mathbf{1}\boldsymbol{A} = -\boldsymbol{A} \\
\boldsymbol{A}^4 &= \boldsymbol{A}^3 \boldsymbol{A} = -\boldsymbol{A}\boldsymbol{A} = \mathbf{1} \\
\boldsymbol{A}^5 &= \boldsymbol{A}
\end{aligned}$$

We can combine these results in the following form:

$$\boldsymbol{A}^n = \begin{cases} (-1)^{\frac{n-1}{2}} \boldsymbol{A} & \text{for odd n} \\ (-1)^{\frac{n}{2}} \mathbf{1} & \text{for even n} \end{cases}$$

Now we insert the powers of $\boldsymbol{A}$ into the power-series expansion:

$$\begin{aligned}
\mathrm{e}^{\boldsymbol{A}t} &= \sum_{k=0}^{\infty} \frac{1}{2k!} (-1)^k \mathbf{1} t^{2k} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} (-1)^k \boldsymbol{A} t^{2k+1} \\
&= \underbrace{\left(\sum_{k=0}^{\infty} \frac{1}{2k!} (-1)^k t^{2k} \right)}_{\cos(t)} \mathbf{1} + \underbrace{\left(\sum_{k=0}^{\infty} \frac{1}{(2k+1)!} (-1)^k t^{2k+1} \right) \boldsymbol{A}}_{\sin(t)}
\end{aligned}$$

We recognize the power-series expansions of the sinus and cosinus functions

$$\begin{aligned}
\cos(x) &= \sum_{k=0}^{\infty} \frac{1}{(2k)!} (-1)^k x^{2k} \\
\sin(x) &= \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} (-1)^k x^{2k+1}
\end{aligned}$$

Thus we obtain our final result

$$\mathrm{e}^{\boldsymbol{A}t} = \cos(t)\mathbf{1} + \sin(t)\boldsymbol{A} \hat{=} \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

Thus $\mathrm{e}^{\boldsymbol{A}t}$ is just the rotation matrix in two dimensions.

I would like to mention here some short-hand notations that are commonly used.

$$\hat{A}|\psi\rangle = |\hat{A}\psi\rangle$$
$$\langle\hat{A}\psi| = \langle\psi|\hat{A}^\dagger$$

We do not use $\langle\psi\hat{A}|$ because that is confusing, if one does not write the hat on the operator, so that one cannot distinguish between operator and state.

5.3 Analogy between states and vectors

Most of what we have learned here is known from vectors. The main difference between states and vectors results from a wave function being complex, while we are used to vectors that are real. The second complication is that states span a space of infinite dimensions, while vectors normally have a finite number of dimensions.

Because wave functions are complex, we needed to introduce bra's and ket's, whereas we use the same symbol for a vector, irrespective of on which side of the scalar product it stands. The vector notation is a specialty of Euclidean spaces. For example when we deal in a curved space, we have to introduce a metric and also then we need to distinguish what are called covariant and contravariant vectors. They are distinguished in the notation by one having a subscript index and the other a superscript index.

A result of the infinite dimension is that the index can be continuous. In that case we have a wave function. The argument of $\Psi(x)$ is nothing more than an index that has become continuous.

When one thinks of a vector one usually has in mind an n-tuple of real numbers. In reality a vector is an abstract object, which has exactly two properties, namely a length and a direction. Describing a vector by an n-tuple of numbers only makes sense after defining a basis. The same is true for states. We usually have in mind wave function in real space, when we think of a state or ket. The state is the abstract object and to imagine it as a function we have introduced a special basis, namely $\{|x\rangle\}$. As we can change the basis and transform into a different coordinate system, we can transform the wave function into a different basis such as momentum eigenstates. The object remained the same, but the basis and hence the numerical representation of the object has changed.

Below, I will show some analogies between vectors and states. Some of the objects described here have not been discussed before, but this will be the subject of the following sections.

Note that the **dyadic** or **outer product** $\vec{s} \otimes \vec{r}$ of two vectors $\vec{s}$ and $\vec{r}$ is a matrix with elements $(\vec{s} \otimes \vec{r})_{i,j} = s_i r_j$. The dyadic product is defined by

$$(\vec{a} \otimes \vec{b})\vec{c} = \vec{a}(\vec{b}\vec{c})$$

The dyadic product must not be mistaken by the vector product or cross product $\vec{r} \times \vec{s}$, which is defined in the three dimensional vector space as a vector.

Abstract object	$\vec{r}$	$\|\psi\rangle$
Inner (scalar) product	$\vec{s}^*\vec{r}$	$\langle\chi\|\psi\rangle$
Outer (dyadic) product	$\vec{r}\otimes\vec{s}^*$	$\|\chi\rangle\langle\psi\|$
Ortho-normality of the basis	$\vec{e}_i\vec{e}_j=\delta_{i,j}$	$\langle\phi_i\|\phi_j\rangle=\delta_{i,j}$
Basis	$\{\vec{e}_i\}$	$\{\|\phi_i\rangle\}$
Representation	$\vec{r}=\sum_i\vec{e}_i r_i$	$\|\psi\rangle=\sum_i\|\phi_i\rangle c_i$
Coefficients	$r_i=\vec{e}_i\vec{r}$	$c_i=\langle\phi_i\|\psi\rangle$
Operator/Matrix	$\vec{s}=\mathbf{A}\vec{r}$	$\|\chi\rangle=\hat{A}\|\psi\rangle$
Operator/Matrix	$s_i=\sum_j A_{i,j}r_j$	$\langle\phi_i\|\chi\rangle=\sum_j\langle\phi_i\|\hat{A}\|\phi_j\rangle\langle\phi_j\|\psi\rangle$

5.4 The power of the Dirac notation

The bracket notation of Dirac makes rather complex quantum mechanical expressions very simple. It seems to me that the beauty is related with the fact that it makes the difference between "lying" and "standing vectors" explicit. Thus we do not need different symbols between scalar and dyadic product as demonstrated in the following equation.

$$\Big(|\Psi_1\rangle\langle\Psi_2|\Big)|\psi_3\rangle=|\Psi_1\rangle\Big(\langle\Psi_2|\psi_3\rangle\Big)$$
$$\Big(\vec{a}_1\otimes\vec{b}\Big)\vec{c}=\vec{a}_1\Big(\vec{b}\vec{c}\Big)$$

In the Dirac notation there is only one product and no parentheses are required.

5.5 Extended Hilbert space

There is an important extension of Hilbert space. This space contains also those elements that can be reached as the limit of a converging series in Hilbert space. It contains states with infinite norm $\langle\psi|\psi\rangle=\infty$. An example for a state that is not an element of Hilbert space, is the δ-function: We can obtain the δ-function as a series of increasingly localized Gaussians. If we consider the norm in this series, it tends to infinity. Another example would be a wave packet with an envelope function that becomes broader until the wave packet becomes a plane wave. If we keep the wave packet normalized, while performing the limit we obtain zero. But if we keep the amplitude at some point in space constant we obtain a plane wave.

The scalar product of a state in this extended Hilbert space with a state in Hilbert space is always finite.

Editorial Remark: Include something about discrete and continuous spectra

5.6 Application: harmonic oscillator

The harmonic oscillator is probably the single most important system in physics. The harmonic oscillator describes a system, for which the Hamilton function is quadratic in the coordinates. In other words, the forces are proportional to the deviation from some equilibrium position. The reason for the importance of the harmonic oscillator is that, at low temperatures, any system will occupy the regions where the potential is the lowest. These are the minima of the potential energy surface. If the potential is smooth, we can approximate it by a Taylor series. If we truncate the Taylor expansion after the first nontrivial term, the result is quadratic. This approximation is a harmonic oscillator.

The harmonic oscillator is, as an example, a model for vibrations of a molecule. Even though a molecule has many degrees of freedom, the model system–a multidimensional harmonic oscillator–can be broken up into decoupled one-dimensional harmonic oscillators.

We discuss the harmonic oscillator at this point, because it illustrates how the abstract notation can be very efficient as compared to wave functions. It will be shown later that one has to rely completely on the abstract notation for certain effects such as the spin of a particle.

5.6.1 Algebraic treatment of the one-dimensional harmonic oscillator

The Hamilton operator for the one dimensional harmonic oscillator has the following form.

$$\hat{H}=\frac{\hat{p}^2}{2m}+\frac{1}{2}c\hat{x}^2=\frac{1}{2m}\left[\hat{p}^2+m^2\omega_0^2\hat{x}^2\right] \tag{5.30}$$

where $\omega_0=\sqrt{\frac{c}{m}}$ is the eigenfrequency of the oscillator defined as $\omega_0=\frac{2\pi}{T}$ by the oscillation period T.

Creation and Annihilation operators

In the following we will show that the the Hamiltonian can be expressed essentially as a product of an operator with its hermitian conjugate. The special property of this operator will allow us to construct the energy spectrum without determining the wave functions.

Let me first sketch the idea of the derivation given below using classical variables: The structure of the Hamiltonian Eq. 5.30 is of the form a^2+b^2. We do know that $a^2-b^2=(a+b)(a-b)$. The minus sign can be converted by replacing $b\to ib$ so that $a^2+b^2=(a+ib)(a-ib)$. If a and b are real we find $a^2+b^2=(a-ib)^*(a-ib)$. What follows is analogous to these steps, only that now we do them with operators instead of numbers.

The Hamilton operator can be written, up to a constant, as a product of two operators.

$$\begin{aligned}
\hat{H} &\overset{\text{Eq. 5.30}}{=} \frac{1}{2m}\left(\hat{p}^2+m^2\omega_0^2\hat{x}^2\right)\\
&= \frac{1}{2m}\Big(\underbrace{(\hat{p}+im\omega_0\hat{x})(\hat{p}-im\omega_0\hat{x})}_{\hat{p}^2-\hat{p}im\omega_0\hat{x}+im\omega_0\hat{x}\hat{p}+m^2\omega_0^2\hat{x}^2}+\underbrace{[\hat{p},im\omega_0\hat{x}]_-}_{+\hat{p}im\omega_0\hat{x}-im\omega_0\hat{x}\hat{p}}\Big)\\
&= \frac{1}{2m}\left[(-i\hat{p}+m\omega_0\hat{x})(i\hat{p}+m\omega_0\hat{x})+m\hbar\omega_0\right]\\
&= \hbar\omega_0\left[\frac{(-i\hat{p}+m\omega_0\hat{x})(i\hat{p}+m\omega_0\hat{x})}{2m\hbar\omega_0}+\frac{1}{2}\right]
\end{aligned}$$

Thus, the Hamiltonian can be written in the form

HAMILTON OPERATOR OF THE HARMONIC OSCILLATOR

$$\hat{H}=\hbar\omega_0\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right) \tag{5.31}$$

where the operators $\hat{a}$ and $\hat{a}^\dagger$ are the creation and annihilation operators[6] for the harmonic oscillator as follows:

[6]The definition of these creation and annihilation operators is not unique. In principle. We may multiply the creation operator with an arbitrary phase factor $e^{i\phi}$ where ϕ is an arbitrary real number. The corresponding annihilation is then mutiplied with the complex conjugate phase factor $e^{-i\phi}$ so that the new pair of operators are again hermitean conjugates of each other. The choice taken here is that the differential operator is real.

CREATION AND ANNIHILATION OPERATORS FOR THE HARMONIC OSCILLATOR

$$\hat{a} = \frac{+i\hat{p} + m\omega_0\hat{x}}{\sqrt{2m\hbar\omega_0}} \tag{5.32}$$

$$\hat{a}^\dagger = \frac{-i\hat{p} + m\omega_0\hat{x}}{\sqrt{2m\hbar\omega_0}} \tag{5.33}$$

We will call the operator a **annihilation operator** and $a^\dagger$ **creation operator**.

We will see below that the operators create and annihilate excitations of the harmonic oscillator. In quantum field theory, the excitations of the harmonic oscillator will be interpreted as particles, so that $a^\dagger$ and a create and annihilate particles. Note that a and $a^\dagger$ are indeed hermitian conjugates of each other.

Spectrum of the harmonic oscillator

In the following we obtain the eigenvalue spectrum of the harmonic oscillator solely from abstract operator equations. We will also find a first order differential equation for the ground state and a recursive equation for the higher states.

Let us evaluate the commutator algebra and let us in addition consider the anti-commutators as well.

ANTI-COMMUTATOR

An **anti-commutator** $[\hat{A}, \hat{B}]_+$ is defined as

$$[\hat{A}, \hat{B}]_+ = \hat{A}\hat{B} + \hat{B}\hat{A} \tag{5.34}$$

It looks like the commutator only with the sign reversed. In order to distinguish the commutator clearly from the anticommutator, a "minus" sign is usually attached to the commutator $[A, B] = [A, B]_- = AB - BA$. Another common notation is to use curly brackets for the anticommutator.

$$\hat{a}^\dagger\hat{a} = \frac{1}{2\hbar\omega_0 m}\hat{p}^2 - \frac{i}{2\hbar}\Big(\hat{p}\hat{x} - \hat{x}\hat{p}\Big) + \frac{m\omega_0}{2\hbar}\hat{x}^2 = \frac{1}{\hbar\omega_0}\hat{H} - \frac{1}{2} \tag{5.35}$$

$$\hat{a}\hat{a}^\dagger = \frac{1}{2\hbar\omega_0 m}\hat{p}^2 + \frac{i}{2\hbar}\underbrace{\Big(\hat{p}\hat{x} - \hat{x}\hat{p}\Big)}_{=[\hat{p},\hat{x}]=\hbar/i} + \frac{m\omega_0}{2\hbar}\hat{x}^2 = \frac{1}{\hbar\omega_0}\hat{H} + \frac{1}{2} \tag{5.36}$$

$$\Big[\hat{a}^\dagger, \hat{a}\Big]_+ \stackrel{\text{Eq. 5.36,5.35}}{=} \frac{2\hat{H}}{\hbar\omega_0} \tag{5.37}$$

$$\Big[\hat{a}^\dagger, \hat{a}\Big]_- \stackrel{\text{Eq. 5.36,5.35}}{=} -\hat{1} \tag{5.38}$$

Let us further consider the commutation relations between $\hat{H}$ and $\hat{a}$, $\hat{a}^\dagger$

$$
\begin{aligned}
\left[\hat{a}, \hat{H}\right]_- &= \left[\hat{a}, \hbar\omega_0(\hat{a}^\dagger\hat{a} + \frac{1}{2})\right]_- \\
&= \hbar\omega_0(\hat{a}\hat{a}^\dagger\hat{a} - \hat{a}^\dagger\hat{a}\hat{a}) = \hbar\omega_0\left[\hat{a}, \hat{a}^\dagger\right]_-\hat{a} \\
&= \hbar\omega_0\hat{a} \\
\left[\hat{a}^\dagger, \hat{H}\right]_- &= \left[\hat{a}^\dagger, \hbar\omega_0(\hat{a}^\dagger\hat{a} + \frac{1}{2})\right]_- \\
&= \hbar\omega_0(\hat{a}^\dagger\hat{a}^\dagger\hat{a} - \hat{a}^\dagger\hat{a}\hat{a}^\dagger) = \hbar\omega_0\hat{a}^\dagger\left[\hat{a}^\dagger, \hat{a}\right]_- \\
&= -\hbar\omega_0\hat{a}^\dagger
\end{aligned}
$$

Thus we found two important relations,

$$
\left[\hat{a}, \hat{H}\right]_- = \hbar\omega_0\hat{a} \qquad \text{and} \qquad \left[\hat{a}^\dagger, \hat{H}\right]_- = -\hbar\omega_0\hat{a}^\dagger, \tag{5.39}
$$

which characterize $\hat{a}^\dagger$ and $\hat{a}$ as creation and annihilation operators. They are a kind of **shift operators** that will play an important role when we discuss symmetries.

The creation and annihilation operators allow us to create new eigenstates from one given eigenstate. If we apply the above equation to an eigenstate of the Hamiltonian $|\psi_n\rangle$ with eigenvalue E_n, we can show that $a|\psi_n\rangle$ is an eigenstate of the Hamiltonian with energy $E_n - \hbar\omega_0$ and that $a^\dagger|\psi_n\rangle$ is and eigenstate with energy $E_n + \hbar\omega_0$.

$$
\begin{aligned}
\hat{H}\hat{a} \quad &= \quad \hat{a}\hat{H} - [\hat{a}, \hat{H}]_- = \hat{a}\Big(\hat{H} - \hbar\omega_0\Big) \\
\Rightarrow \hat{H}\Big(\hat{a}|\psi_n\rangle\Big) \quad &\overset{\hat{H}|\psi_n\rangle = |\psi_n\rangle E_n}{=} \quad \Big(\hat{a}|\psi_n\rangle\Big)\Big(E_n - \hbar\omega_0\Big) \qquad (5.40) \\
\hat{H}\hat{a}^\dagger \quad &= \quad \hat{a}^\dagger\hat{H} - [\hat{a}^\dagger, \hat{H}]_- = \hat{a}^\dagger(\hat{H} + \hbar\omega_0) \\
\Rightarrow \hat{H}\Big(\hat{a}^\dagger|\psi_n\rangle\Big) \quad &\overset{\hat{H}|\psi_n\rangle = |\psi_n\rangle E_n}{=} \quad \Big(\hat{a}^\dagger|\psi_n\rangle\Big)(E_n + \hbar\omega_0) \qquad (5.41)
\end{aligned}
$$

Thus we can create a seemingly infinite number of eigenstates $|\Psi_{n-k}\rangle = a^k|\psi_n\rangle$ with equi-spaced eigenvalues $E_{n-k} = E_n - k\hbar\omega_0$. Is there an end to it? Yes, the spectrum is bound below and therefore the series must end. The spectrum is bound, because the expectation value of the Hamiltonian can be written as the norm of a state, which is always nonnegative and a constant, namely

$$
\begin{aligned}
\langle\psi|\hat{H}|\psi\rangle &= \langle\psi|\hbar\omega_0\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)|\psi\rangle \\
\Rightarrow \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} &= \hbar\omega_0\left(\frac{\langle\psi|\hat{a}^\dagger\hat{a}|\psi\rangle}{\langle\psi|\psi\rangle} + \frac{1}{2}\right) = \hbar\omega_0\left(\frac{\langle\hat{a}\psi|\hat{a}\psi\rangle}{\langle\psi|\psi\rangle} + \frac{1}{2}\right) \geq 0.
\end{aligned}
$$

We denote the ground state by $|\psi_0\rangle$. We have assumed that the states are normalized.

The ground state fulfills

$$
\hat{a}|\psi_0\rangle = 0 \tag{5.42}
$$

because otherwise there would be another eigenstate $a|\psi_0\rangle$ with lower energy.

What is the energy of this ground state?

$$
\hat{H}|\psi_0\rangle \overset{\text{Eq. 5.31}}{=} \hbar\omega_0\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)|\psi_0\rangle \overset{\text{Eq. 5.42}}{=} \frac{1}{2}\hbar\omega_0|\psi_0\rangle
$$

The energy of the ground state is $E_0 = \frac{1}{2}\hbar\omega_0$.

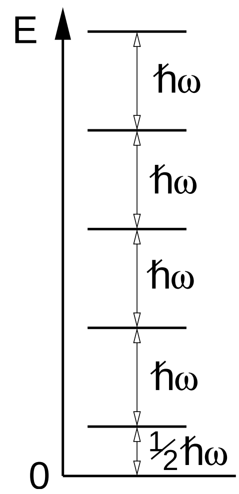
ENERGY SPECTRUM OF THE HARMONIC OSCILLATOR

The energy eigenvalue spectrum of the harmonic oscillator is equi-spaced

$$E_n = \hbar\omega_0(n + \frac{1}{2}) \tag{5.43}$$

where n is a non-negative integer.

Again we observe that the ground state energy does not vanish, just as for the particle in a box. The ground state energy is due to the so-called **zero-point vibrations** or **zero-point motion**. It is not possible to remove that remaining energy out of the system. The origin of the zero-point vibrations is Heisenberg's uncertainty relation , which says that it is not possible to find quantum mechanical system with a determined position and momentum. That would be required for the classical system to be at rest. Can we observe the zero point vibration? We cannot observe it by removing that energy, however if we compare the energy of an hydrogen molecule and a deuterium molecule, which have identical forces acting between the two atoms. (Deuterium is an isotope of hydrogen with a mass about two times larger.) We find that the deuterium molecule is more stable by a few tenths of an eV which is entirely due to the different zero-point energies.



5.6.2 Wave functions of the harmonic oscillator

From the ground state all excited states can be constructed by repeatedly applying the creation operator. However, we need to add normalization factor so that all states are properly normalized.

$$\begin{aligned}
\langle \hat{a}^\dagger \psi_n | \hat{a}^\dagger \psi_n \rangle &= \langle \psi_n | \hat{a}\hat{a}^\dagger | \psi_n \rangle \stackrel{\text{Eq. 5.36}}{=} \langle \psi_n | \frac{\hat{H}}{\hbar\omega} + \frac{1}{2} | \psi_n \rangle \\
&\stackrel{\text{Eq. 5.43}}{=} \langle \psi_n | \psi_n \rangle (n+1) \\
\Rightarrow \qquad |\psi_{n+1}\rangle &= \frac{1}{\sqrt{n+1}} \hat{a}^\dagger |\psi_n\rangle & (5.44) \\
\langle \hat{a}\psi_n | \hat{a}\psi_n \rangle &= \langle \psi_n | \hat{a}^\dagger \hat{a} | \psi_n \rangle \stackrel{\text{Eq. 5.35}}{=} \langle \psi_n | \frac{\hat{H}}{\hbar\omega} - \frac{1}{2} | \psi_n \rangle \\
&\stackrel{\text{Eq. 5.43}}{=} \langle \psi_n | \psi_n \rangle n \\
\Rightarrow \qquad |\psi_{n-1}\rangle &= \frac{1}{\sqrt{n}} \hat{a} |\psi_n\rangle & (5.45)
\end{aligned}$$

EIGENSTATES OF THE HARMONIC OSCILLATOR

The eigenstates of the harmonic oscillator can be constructed by successively applying the creation operator to the ground state $|\psi_0\rangle$.

$$|\psi_n\rangle \stackrel{\text{Eq. 5.44}}{=} \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n|\psi_0\rangle \tag{5.46}$$

The ground state is defined by Eq. 5.42.

$$\hat{a}|\psi_0\rangle = 0 \tag{5.47}$$

What is the wave function of the ground state? We can obtain it from

$$\hat{a}|\psi_0\rangle = 0 \tag{5.48}$$

which is, with Eq. 5.32, a first order differential equation

$$\Big(\hbar\partial_x + m\omega_0 x\Big)\psi_0(x) = 0$$

A first order differential equation has a tremendous advantage over a second order differential equation because it can be mapped onto a simple integration using the Ansatz

$$\psi_0(x) = \mathrm{e}^{S(x)}$$

With this Ansatz we obtain

$$\begin{aligned}
\hbar\partial_x S + m\omega_0 x &= 0\\
S(x) &= \int_{-\infty}^{x} dx'\Big(-\frac{m\omega_0}{\hbar}x'\Big) = C - \frac{m\omega_0}{2\hbar}x^2\\
\psi_0(x) &= \mathrm{e}^{S(x)} = A\mathrm{e}^{-\frac{m\omega_0}{2\hbar}x^2} = \sqrt[4]{\frac{m\omega_0}{\pi\hbar}}\mathrm{e}^{-\frac{m\omega_0}{2\hbar}x^2}
\end{aligned}$$

where $A = \mathrm{e}^C$ is the normalization constant. The ground state wave function of the harmonic oscillator is a simple Gaussian.

The higher eigenstates are obtained by applying $a^\dagger$ several times to the ground state. This is even simpler that finding the ground state, because we only need to form the derivatives of the already known states. The result are a polynomials times the Gaussian. The polynomials are the so-called **Hermite polynomials**. Some useful relations about Hermite polynomials can be found in the appendix of "Quantum Mechanics" Vol. I by A. Messiah.

5.6.3 Multidimensional harmonic oscillator

The multidimensional harmonic operator describes for example the vibrations of a molecule. The vibrations can be tested using infrared (IR) spectroscopy, where the absorption of light is determined as function of frequency. Light is absorbed when a vibration is excited from the ground state to an excited state. The excitation energy taken from a photon that is absorbed and a phonon (or lattice vibration) is "emitted". As the vibrational energies are characteristic of the atomic structure, IR spectroscopy is an important tool to understand the nature of the material under study. The atomic vibrations can be described as a multidimensional harmonic oscillator. To lowest order the atoms oscillate about their equilibrium positions. We will see that higher order terms of the Taylor-expansion of the potential about the equilibrium structure results in an interaction between vibrations, so that, in a crystal, we talk about phonon-phonon interactions. For further information on IR spectroscopy see Atkins[4]

The multidimensional oscillator has the Hamiltonian

$$H = \sum_i \frac{\hat{p}_i^2}{2m_i} + \frac{1}{2}\sum_{i,j} \hat{x}_i c_{i,j} \hat{x}_j$$

It is always possible to transform this Hamiltonian onto a system of independent one-dimensional harmonic oscillators shown as follows.

First we introduce new coordinates $\hat{X}_i = \sum_j A_{i,j}\hat{x}_j$ with a real matrix $\boldsymbol{A}$.[7] Per definition, the canonical conjugate momentum $\hat{P}$ of $\hat{X}$ must fulfill the commutator relation[8] [9] $[\hat{P}, \hat{X}] = \frac{\hbar}{i}$, which determines the momentum as $\hat{P}_i = \sum_j A^{-1,\top}_{i,j}\hat{p}_j$. Because we required the matrix $\boldsymbol{A}$ to be real, we can replace $\boldsymbol{A}^{-1,\top}$ by $\boldsymbol{A}^{-1,\dagger}$.

TRANSFORMATION OF MOMENTA UNDER A COORDINATE TRANSFORM

$$\hat{x}'_i = \sum_j A_{i,j} x_j \qquad \Leftrightarrow \qquad \hat{p}'_i = \sum_j A^{-1,\dagger}_{i,j} p_j \tag{5.49}$$

The transformation matrix $\boldsymbol{A}$ is real. The transformation of momenta is chose such that the transformed coordinates and momenta obey the commutator relation $[\hat{p}'_i, \hat{x}'_j] = \frac{\hbar}{i}\delta_{i,j}$

We will use two such transformations in the following derivation. The first transformation is a rescaling of the positions and momenta to get rid of the masses.

$$\hat{\xi}_i \stackrel{\text{def}}{=} \sqrt{m_i}\hat{x}_i$$
$$\hat{\pi}_i \stackrel{\text{def}}{=} \frac{1}{\sqrt{m_i}}\hat{p}_i$$

so that the Hamiltonian has the form

$$H = \sum_i \frac{1}{2}\hat{\pi}_i^2 + \frac{1}{2}\sum_{i,j} \hat{\xi}_i \underbrace{\frac{1}{\sqrt{m_i}} c_{i,j} \frac{1}{\sqrt{m_j}}}_{D_{i,j}} \hat{\xi}_j$$

[7] The generalization to complex transformations is at least problematic.

[8] The transformation rules for the momenta can be obtained from classical physics as follows: Let us consider a transformation of the coordinates.

$$x_i = f_i(\vec{x}', t) \qquad \Rightarrow \qquad \dot{x}_i = \sum_j \frac{\partial f_i}{\partial x'_j}\dot{x}'_j + \frac{\partial f_i}{\partial t}$$

We already calculated the transformation for the velocities, which will be required in the following

Now we express the action by the old and the new coordinates

$$S[x(t)] = \int dt\; \mathcal{L}(\vec{x}, \dot{\vec{x}}, t) = \int dt\; \mathcal{L}(\vec{f}(\vec{x}', t), \sum_j \frac{\partial \vec{f}}{\partial x'_j}\dot{x}'_j + \frac{\partial \vec{f}}{\partial t}, t) = \int dt\; \mathcal{L}'(\vec{x}', \dot{\vec{x}}', t) = S[x'(t)]$$

The momentum is defined related to the old and the new (primed) coordinates is

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}_i}; \qquad \text{and} \qquad p'_i = \frac{\partial \mathcal{L}'}{\partial \dot{x}'_i} = \sum_j \frac{\partial \mathcal{L}}{\partial \dot{x}_j}\frac{\partial f_j}{\partial x'_i} = \sum_j p_j \frac{\partial f_j}{\partial x'_i}$$

This provides the transformation rule for momenta for a given transformation of positions.

[9]

$$[\hat{P}_i, \hat{X}_j] = \sum_{m,n}[A^{-1}_{m,j}\hat{p}_m, A_{i,n}\hat{x}_n] = \sum_{m,n} A^{-1}_{m,j}A_{i,n}\underbrace{[\hat{p}_m, \hat{x}_n]}_{\frac{\hbar}{i}\delta_{m,n}} = \frac{\hbar}{i}\underbrace{\sum_m A_{i,m}A^{-1}_{m,j}}_{\delta_{i,j}} = \frac{\hbar}{i}$$

The matrix $D_{i,j} \stackrel{\text{def}}{=} \frac{1}{\sqrt{m_i}} c_{i,j} \frac{1}{\sqrt{m_j}}$ is called **dynamical matrix**.

We diagonalize the dynamical matrix so that

$$\sum_k D_{i,k} U_{k,j} = U_{i,j} d_j$$

the vector $\{U_{k,j}\}$ for a given j is an eigenvector of $\boldsymbol{D}$ and d_j is the corresponding eigenvalue. We assume that the eigenvectors are normalized. Since the eigenvectors of a Hermitian matrix such as $\boldsymbol{D}$ are orthonormal, $\boldsymbol{U}$ is a unitary matrix.

$$\sum_k U^*_{k,i} U_{k,j} = \delta_{i,j}$$

Especially for real matrices, the eigenvectors can be chosen to be real.[10] This corresponds to a transformation, where the transformed coordinate axes point along the principal axes.[11]

Now we are ready to do the second transformation[12]

$$\hat{X}_i \stackrel{\text{def}}{=} \sum_j U^\dagger_{i,j} \hat{\xi}_j = \sum_j U^\dagger_{i,j} \sqrt{m_j} \hat{x}_j$$
$$\hat{P}_i \stackrel{\text{def}}{=} \sum_j U^\dagger_{i,j} \hat{\pi}_j = \sum_j U^\dagger_{i,j} \frac{1}{\sqrt{m_j}} \hat{p}_j$$

which yields[13]

$$H = \sum_i \Big(\frac{1}{2} \hat{P}_i^2 + \frac{1}{2} d_i \hat{X}_i^2 \Big)$$

Finally we use the method of separation of variables: We use the product Ansatz $\Psi(X_1, \ldots, X_n) = \prod_{i=1}^n \phi_i(X_i)$, where each wave function $\phi_i(X)$ fulfills the Schrödinger equation for the one-dimensional harmonic oscillator

$$\hat{H}_i |\phi_i\rangle = |\phi_i\rangle e_i$$
$$\hat{H}_i = \frac{1}{2} \hat{P}_i^2 + \frac{1}{2} d_i \hat{X}_i^2$$

There are of course many eigenvalues and eigenstates for each one-dimensional harmonic oscillator. For reasons of simplicity we suppress the corresponding index.

The eigenvalues for the multidimensional harmonic oscillator are

$$E = \sum_i e_i$$

and the wave functions are obtained as

$$\Psi(x_1, \ldots, x_n) = \prod_{i=1}^n \phi_i \Big(\sum_k U_{ik} \sqrt{m_k} x_k \Big)$$

[10] The eigenvectors are not necessarily real,which is important to know when using canned tools for the diagonalization. For real, hermitian matrices, however, the eigenvectors can be chosen real, because otherwise the form $D_{i,j} = \sum_k U_{ik} d_k U^*_{j,k}$ would not be real.

[11] germ(principal axis) = Hauptachse

[12] Note that $(\boldsymbol{U}^\dagger)^{-1\dagger} = \boldsymbol{U}^\dagger$ since $\boldsymbol{U}$ is unitary.

[13]

$$\sum_i \hat{\pi}_i^2 = \sum_{i,j,k} U_{i,j} U_{i,k} \hat{P}_j \hat{P}_k \stackrel{\boldsymbol{U}=\boldsymbol{U}^*}{=} \sum_{i,j,k} U^*_{i,j} U_{i,k} \hat{P}_j \hat{P}_k = \sum_{j,k} \hat{P}_j \underbrace{\Big(\sum_i U^\dagger_{j,i} U_{i,k} \Big)}_{\delta_{j,k}} \hat{P}_k = \sum_j \hat{P}_j^2$$

Chapter 6

Representations

In the last section we introduced abstract states and operators that transform them into each other. The only way to produce numbers is to evaluate scalar products and matrix elements. If we wish to represent a state numerically, for example as wave function, we need to choose a representation. Choosing a representation in quantum mechanics is like choosing a coordinate system for vectors. In quantum mechanical calculations, the choice of representations and the transformation between representations is very important.

In this section we will learn about the matrix formulation of quantum mechanics, where a state is represented by a vector and an operator is represented by a matrix. The only difficulty is that we need to deal with infinitely many dimensions.

We will also see the connection to wave functions as used in Schrödinger's wave-mechanical formulation of quantum mechanics, where states are functions in space. The wave function can be thought of as a vector with a continuous vector index.

Fig. 6.1: Werner Heisenberg, 1901-1976. German Physicist. Nobel price 1933 in physics for the invention of Quantum Mechanics, which he published only 23 years old.

6.1 Unity operator

Let us consider a complete set of states $\{|u_i\rangle\}$. Then we can express the **unity operator** as

$$\hat{1} = \sum_{i,j} |u_i\rangle K_{i,j} \langle u_j| \tag{6.1}$$

where $\boldsymbol{K}$ is the inverse of the so-called overlap matrix $\boldsymbol{O}$ defined as $O_{i,j} = \langle u_i | u_j \rangle$.

$$\sum_k K_{i,k} \langle u_k | u_j \rangle = \delta_{i,j} \tag{6.2}$$

PROOF: we show that every state is mapped onto itself, if the unity operator in this form is applied to it. Every state $|\psi\rangle$ can be expressed as $|\psi\rangle = \sum_i |u_i\rangle c_i$ if $\{|u_i\rangle\}$ is complete.

$$\Big(\sum_{i,j} |u_i\rangle K_{i,j}\langle u_j|\Big)|\psi\rangle = \Big(\sum_{i,j} |u_i\rangle K_{i,j}\langle u_j|\Big)\Big(\sum_k |u_k\rangle c_k\Big)$$
$$= \sum_{i,k} |u_i\rangle \underbrace{\Big(\sum_j K_{i,j}\langle u_j|u_k\rangle\Big)}_{=\delta_{i,k} \quad \text{Eq. 6.2}} c_k = \sum_i |u_i\rangle c_i = |\psi\rangle$$

q.e.d.

If the set $\{|u_i\rangle\}$ is not only complete but also orthonormal, the overlap matrix $\boldsymbol{O}$ and its inverse $\boldsymbol{K}$ are unity matrices. Hence,

$$\hat{1} = \sum_i |u_i\rangle\langle u_i|$$

for an orthonormal basis $\{u_i\}$. In the following we assume $\{|u_i\rangle\}$ to be orthonormal.

6.2 Representation of a state

We can now express a state $|\psi\rangle$ in a certain **representation**. A representation is defined by a complete basis of orthonormal states $\{|u_i\rangle\}$. A state can be represented in a certain basis by multiplication with the unity operator $\sum_i |u_i\rangle\langle u_i|$.

$$|\psi\rangle = \sum_i |u_i\rangle\langle u_i|\psi\rangle$$

Thus we can represent the state $|\psi\rangle$ by a vector $\vec{\psi}$, where each vector component is a number, namely $\psi_i = \langle u_i|\psi\rangle$. However, we need to specify the **basis**, namely the set $\{|u_i\rangle\}$. Compared to ordinary vector algebra, the only uncommon points is that the vector has infinitely many components and that the vector components are complex numbers.

What we have done is similar to representing a vector $\vec{r}$ in three dimensions in a basis spanned by three orthonormal basis vectors $\vec{e}_x, \vec{e}_y, \vec{e}_z$. We can write

$$\vec{r} = \vec{e}_x r_x + \vec{e}_y r_y + \vec{e}_z r_z$$

Usually the vector $\vec{r}$ is then represented by its three components (r_x, r_y, r_z). In order to know the vector $\vec{r}$ we need to know the choice of basis vectors $\vec{e}_x, \vec{e}_y, \vec{e}_z$ and the vector components r_x, r_y, r_z in this basis.

A special representation is the real space representation. Here the basis $\{|r\rangle\}$ is continuous. The scalar products between states of the basis are defined as $\langle r|r'\rangle = \delta(r - r')$. Instead of a sum we need to choose the integral. The unity operator has the form

$$\hat{1} = \int d^3r |r\rangle\langle r|$$

Thus we obtain

$$|\psi\rangle = \int d^3r |r\rangle\langle r|\psi\rangle = \int d^3r |r\rangle\psi(r)$$

so that we can represent the state by a function, the wave function $\psi(r) = \langle r|\psi\rangle$.

6.3 Representation of an operator

Let us now represent an operator A in the basis

$$\hat{A} = \Big(\sum_i |u_i\rangle\langle u_i|\Big)\hat{A}\Big(\sum_j |u_j\rangle\langle u_j|\Big)$$
$$= \sum_{i,j} |u_i\rangle\langle u_i|\hat{A}|u_j\rangle\langle u_j|$$

The operator A in the representation is a matrix $\mathbf{A}$ with matrix elements $A_{i,j} = \langle u_i|A|u_j\rangle$.

If we consider a general matrix element of A we can write it in Bracket or vector representation

$$\langle\phi|\hat{A}|\psi\rangle = \sum_{i,j}\langle\phi|u_i\rangle\langle u_i|\hat{A}|u_j\rangle\langle u_j|\psi\rangle = \sum_{i,j}\phi_i^* A_{i,j}\psi_j$$

as vector-matrix-vector product. Thus a general matrix element can be written in a given representation as vector-matrix-vector product.

6.4 Change of representations

It is convenient to choose the representation according to the problem at hand. Therefore, it is described here how to get from one representation to the other.

Let us start from a representation $\{|u_i\rangle\}$, and pass over to a representation defined by the base $\{|v_i\rangle\}$. In order to discriminate between the representations we place the symbol of the representation in parenthesis as superscript, when necessary.

$$|\psi\rangle = \underbrace{\sum_i |v_i\rangle\langle v_i|}_{=\hat{1}}\underbrace{\sum_j |u_j\rangle\langle u_j|}_{=\hat{1}}|\psi\rangle = \sum_i |v_i\rangle \underbrace{\sum_j \overbrace{\langle v_i|u_j\rangle}^{U^\dagger_{i,j}}\overbrace{\langle u_j|\psi\rangle}^{\psi_j^{(u)}}}_{\psi_i^{(v)}}$$

Thus the state vector $\vec{\psi}$ is transformed by multiplication with a matrix $\boldsymbol{U}^\dagger$,

$$\psi_i^{(v)} = \sum_j U^\dagger_{i,j}\psi_j^{(u)}$$
$$U_{i,j} = \langle u_i|v_j\rangle$$

The meaning of the matrix $\boldsymbol{U}$ is that it transforms the orthonormal basis $\{|u_i\rangle\}$ into the other orthonormal basis $\{|v_i\rangle\}$, that is

$$|v_i\rangle = \sum_j |u_j\rangle\underbrace{\langle u_j|v_i\rangle}_{U_{j,i}} = \sum_j |u_j\rangle U_{j,i}$$

The matrix $\boldsymbol{U}$, which describes the transformation between two orthonormal basissets is a unitary matrix.
PROOF: The matrix U is defined as $U_{i,j} = \langle u_i|v_j\rangle$, where $\{|u_i\rangle\}$ and $\{|v_i\rangle\}$ are orthonormal and complete.

$$\sum_k U^\dagger_{i,k}U_{k,j} = \sum_k\langle v_i|u_k\rangle\langle u_k|v_j\rangle = \langle v_i|\Big(\sum_k |u_k\rangle\langle u_k|\Big)|v_j\rangle = \langle v_i|v_j\rangle = \delta_{i,j}$$

q.e.d.

The transformation of matrix elements proceeds similar and we obtain

$$
\begin{aligned}
\langle v_i|\hat{A}|v_j\rangle &= \langle v_i|\Big(\sum_k |u_k\rangle\langle u_k|\Big)\hat{A}\Big(\sum_l |u_l\rangle\langle u_l|\Big)|v_j\rangle \\
A^{(v)}_{i,j} &= \sum_{k,l} U^\dagger_{i,k}A^{(u)}_{k,l}U_{l,j}
\end{aligned}
$$

Let us summarize the transformation rules:

CHANGE OF REPRESENTATION

Under a change from one orthonormal basis to another, the vector components of a state and the matrix elements transform as

$$
\psi^{(v)}_i = \sum_j U^\dagger_{i,j}\psi^{(u)}_j \tag{6.3}
$$

$$
A^{(v)}_{i,j} = \sum_{k,l} U^\dagger_{i,k}A^{(u)}_{k,l}U_{l,j} \tag{6.4}
$$

where

$$
U_{i,j} = \langle u_i|v_j\rangle \tag{6.5}
$$

is a unitary matrix.

6.5 From bra's and ket's to wave functions

6.5.1 Real-space representation

Now we have learned to deal with abstract states and with wave functions. What is missing is how to translate the expressions from one formulation to the other. While the connection is better shown in the context of representation theory discussed later in chapter 6, I will show here already the main arguments.

The connection between the two formulations lies in

$$
\Psi(x)\stackrel{\text{def}}{=}\langle x|\Psi\rangle \tag{6.6}
$$

where the states $|x\rangle$ form a complete set of states. Note, however, that the states $|x\rangle$ are not elements of the Hilbert space, but lie in the extended Hilbert space.

Similarly, the meaning of of the somewhat sloppy expression $\hat{A}\psi(x)$, where $\hat{A}$ is an arbitrary operator, is as follows

$$
\hat{A}\psi(x)\stackrel{\text{def}}{=}\langle x|\hat{A}|\psi\rangle
$$

The states $|x\rangle$ obey the conditions:

$$
\hat{x}|x\rangle = |x\rangle x \qquad \text{eigenvalue equation} \tag{6.7}
$$

$$
\langle x|x'\rangle = \delta(x-x') \qquad \text{orthonormality} \tag{6.8}
$$

$$
\hat{1} = \int dx\, |x\rangle\langle x| \qquad \text{completeness} \tag{6.9}
$$

The second equation, Eq. 6.8, shows that the wave function related to the state $|x_0\rangle$, defined according to Eq. 6.6, is a δ-function located at position x_0.

Position and momentum operators have the form[1]

$$\hat{x} = \int dx |x\rangle x \langle x| \tag{6.10}$$
$$\hat{p} = \int dx |x\rangle \frac{\hbar}{i}\partial_x \langle x| \tag{6.11}$$

The Schrödinger equation has the following form in the abstract notation and the wave function notation.

$$\left(\frac{\hat{p}^2}{2m} + v(\hat{x})\right)|\psi\rangle = |\Psi\rangle E \tag{6.12}$$
$$\left(\frac{-\hbar^2}{2m}\partial_x^2 + v(x)\right)\psi(x) = \psi(x)E \tag{6.13}$$

The two above versions of the Schrödinger equations are mathematically correct. However, an equation like the following is, strictly speaking, meaningless

$$\left[\frac{-\hbar^2}{2m}\partial_x^2 + v(x)\right]|\psi\rangle = |\psi\rangle E$$

because a function in space is multiplied with a abstract state. Nevertheless, such a notation is often used in practice.

The difficult steps in the derivation of the above expression Eq. 6.13 from Eq. 6.12 are

1. to represent a function of an operator

$$\begin{aligned} v(\hat{x})|x\rangle &\overset{\text{Eq. 5.29}}{=} |x\rangle v(x) \\ \Rightarrow v(\hat{x}) = v(\hat{x})\underbrace{\int dx|x\rangle\langle x|}_{\hat{1}} &= \int dx\, |x\rangle v(x)\langle x| \end{aligned}$$

[1]We first show that the first equation Eq. 6.10 holds:

$$\hat{x} \overset{\text{Eq. 6.9}}{=} \hat{x}\underbrace{\int dx\, |x\rangle\langle x|}_{\hat{1}} = \int dx\, \underbrace{\hat{x}|x\rangle}_{|x\rangle x}\langle x| \overset{\text{Eq. 6.7}}{=} \int dx\, |x\rangle x\langle x|$$

Now we show that the second equation Eq. 6.11 holds: The momentum operator is defined as that operator that has a commutator $[\hat{p},\hat{x}] = \frac{\hbar}{i}$. Thus we need to show that the commutator has the correct value.

$$\begin{aligned} [\hat{p},\hat{x}] &\overset{\text{Eq. 5.25}}{=} \hat{p}\hat{x} - \hat{x}\hat{p} \\ &\overset{\text{Eq. 6.11}}{=} \int dx|x\rangle\frac{\hbar}{i}\partial_x\langle x|\int dx'\, |x'\rangle x'\langle x'| - \int dx'\, |x'\rangle x'\langle x'|\int dx|x\rangle\frac{\hbar}{i}\partial_x\langle x| \\ &= \int dx\int dx' \left(|x\rangle\frac{\hbar}{i}\partial_x \underbrace{\langle x|x'\rangle}_{\delta(x-x')} x'\langle x'| - |x'\rangle x' \underbrace{\langle x'|x\rangle}_{\delta(x-x')} \frac{\hbar}{i}\partial_x\langle x|\right) \\ &= \int dx\, \left(|x\rangle\frac{\hbar}{i}\partial_x x\langle x| - |x\rangle x\frac{\hbar}{i}\partial_x\langle x|\right) \\ &= \int dx\, |x\rangle\Big(\underbrace{\frac{\hbar}{i}\partial_x x}_{\frac{\hbar}{i}+x\frac{\hbar}{i}\partial_x} - x\frac{\hbar}{i}\partial_x\Big)\langle x| = \frac{\hbar}{i} \end{aligned}$$

2. to square the momentum operator

$$
\begin{aligned}
\hat{p}^2 &\stackrel{\text{Eq. 6.11}}{=} \left(\int dx\, |x\rangle \frac{\hbar}{i}\partial_x \langle x|\right)\left(\int dx'\, |x'\rangle \frac{\hbar}{i}\partial_{x'}\langle x'|\right)\\
&= \int dx\, |x\rangle \frac{\hbar}{i}\partial_x \left(\int dx'\, \underbrace{\langle x|x'\rangle}_{\delta(x-x')} \frac{\hbar}{i}\partial_{x'}\langle x'|\right)\\
&= \int dx\, |x\rangle \frac{\hbar}{i}\partial_x \left(\frac{\hbar}{i}\partial_x \langle x|\right)\\
&= \int dx\, |x\rangle(-\hbar^2\partial_x^2)\langle x|
\end{aligned}
$$

6.5.2 Momentum representation

What are the momentum eigenstates? We start from the eigenvalue equation for the momentum operator, which defines the momentum eigenstates $|p\rangle$ up to a constant.

$$\hat{p}|p\rangle = |p\rangle p \qquad \text{eigenvalue equation} \tag{6.14}$$

We would like to know the momentum eigenstates as wave functions: With the help of Eq. 6.11 we obtain an eigenvalue equation for $\langle x|p\rangle$.

$$\int dx\, |x\rangle \frac{\hbar}{i}\partial_x \langle x|p\rangle \stackrel{\text{Eq. 6.14}}{=} \int dx\, |x\rangle\langle x|p\rangle p$$

Now, we exploit that the kets $\{|x\rangle\}$ are linear independent. This implies that the coefficients for the kets $|x\rangle$ must be pairwise identical.

$$\Rightarrow \frac{\hbar}{i}\partial_x \langle x|p\rangle = \langle x|p\rangle p$$

This is a differential equation with constant coefficients for the function $f(x) \stackrel{\text{def}}{=} \langle x|p\rangle$, which can be solved with the exponential ansatz. As solution, one obtains a simple plane wave

$$\langle x|p\rangle = \mathrm{e}^{\frac{i}{\hbar}px} \tag{6.15}$$

One should be careful not to write $\langle x|p\rangle = p(x)$ even though this is formally correct: One could easily mistake it for a function that expresses the momentum as function of position.

Note that the prefactor of the wave function can be chosen in different ways. In our notation the orthonormality condition and the representation of the unity operator are.

ORTHONORMALITY AND COMPLETENESS OF MOMENTUM EIGENSTATES

$$\langle p|p'\rangle = 2\pi\hbar\, \delta(p-p') \qquad \text{orthonormality} \tag{6.16}$$

$$\hat{1} = \int \frac{dp}{2\pi\hbar}\, |p\rangle\langle p| \qquad \text{completeness} \tag{6.17}$$

In many textbooks, the momentum eigenstates are defined as $|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\mathrm{e}^{\frac{i}{\hbar}px}$. That is they are normalized differently. Our notation has the advantage that the factors $\frac{1}{\sqrt{2\pi\hbar}}$ do not appear in the real space wave function corresponding to the momentum eigenstate. On the other hand, our notation requires the factors $2\pi\hbar$ in the normalization condition Eq. 6.16 and the representation of the unit operator, Eq. 6.17, which do not occur in the other notation.

6.5.3 Orthonormality condition of momentum eigenstates (Home study)

Here, we derive the normalization condition of the momentum eigenstates, namely Eq. 6.16. We use the expression Eq. 6.15 for the real space wave function of a momentum eigenstate.

$$\langle p|p'\rangle = \langle p|\overbrace{\int_{-\infty}^{\infty} dx\,|x\rangle\langle x|}^{\hat{1}}|p'\rangle = \int_{-\infty}^{\infty} dx\,\langle p|x\rangle\langle x|p'\rangle = \int_{-\infty}^{\infty} dx\,\mathrm{e}^{\frac{i}{\hbar}(p-p')x} \tag{6.18}$$

This expression is awkward, because it contains an integral that does not even converge, if the bounds of the integral are at infinity. Therefore, we have to play a common trick to make the integral converge anyway[2]: We multiply the integrand with a function, namely $\mathrm{e}^{-\lambda|x|}$, that attenuates the oscillations for large arguments, i.e. at $\pm\infty$. For each value of λ, the integral is well defined. As λ approaches zero, the function $\mathrm{e}^{-\lambda|x|}$ approaches 1. Hence, we recover the original integral Eq. 6.18 in the limit $\lambda \to 0$.

$$\begin{aligned}
\langle p|p'\rangle &= \lim_{\lambda\to 0^+}\int_{-\infty}^{\infty} dx\,\mathrm{e}^{(\frac{i}{\hbar}(p-p')x}\mathrm{e}^{-\lambda|x|} \\
&= \lim_{\lambda\to 0}\left(\int_{-\infty}^{0} dx\,\mathrm{e}^{(\frac{i}{\hbar}(p-p')+\lambda)x} + \int_{0}^{\infty} dx\,\mathrm{e}^{(\frac{i}{\hbar}(p-p')-\lambda)x}\right) \\
&= \lim_{\lambda\to 0}\left(\left(\frac{i}{\hbar}(p-p')+\lambda\right)^{-1} - \left(\frac{i}{\hbar}(p-p')-\lambda\right)^{-1}\right) \\
&= \lim_{\lambda\to 0}\frac{-2\lambda}{\frac{-1}{\hbar^2}(p-p')^2-\lambda^2} \\
&= 2\hbar\lim_{\lambda\to 0}\frac{\hbar\lambda}{(p-p')^2+\hbar^2\lambda^2}
\end{aligned} \tag{6.19}$$

The function in the limes has the form of a **Lorentzian**. A Lorentzian is defined as $L(x) = \frac{1}{\pi}\frac{\Gamma/2}{(x-x_0)^2+(\Gamma/2)^2}$. The Lorentzian has a peak at x_0 and a width related to Γ. The Lorentzian has the form of a spectral line in an absorption spectrum.

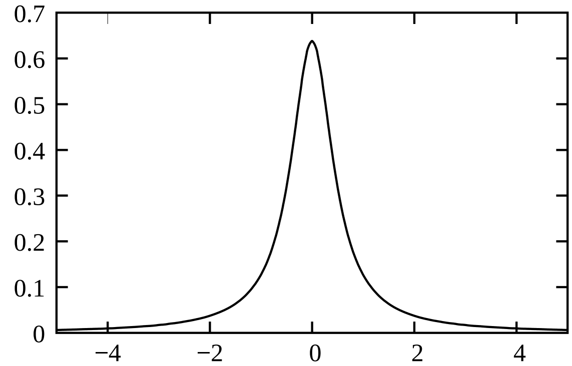


Fig. 6.2: Lorentzian with width $\Gamma = 1$

We recognize already that the function in Eq. 6.19 vanishes in the limit for $p \neq p'$ and that it diverges for $p = p'$, which suggests that the limes could be a delta function. What needs to be shown is that the integral over p is equal to a well-defined constant.

[2]The background is the theory of Distributions, which are generalizations of functions such as the delta function of Dirac. See [25]. **Editor: In the first line we interchange the limit and the integral. Is this legitimate?**

We use the integral formula taken from Bronstein[26]

$$\int dx \frac{1}{x^2+a^2} = \frac{1}{a}\mathrm{atan}(\frac{x}{a})$$

Thus we obtain

$$\begin{aligned}\int_{-\infty}^{\infty} dp\, \langle p|p'\rangle &\stackrel{\text{Eq. 6.19}}{=} 2\hbar \lim_{\lambda\to 0}\left[\mathrm{atan}(\frac{p-p'}{\hbar\lambda})\right]_{-\infty}^{\infty} = 2\hbar\left[\underbrace{\mathrm{atan}(\infty)}_{\pi/2} - \underbrace{\mathrm{atan}(-\infty)}_{-\pi/2}\right] \\ &= 2\pi\hbar\end{aligned}$$

Since a function that integrates to one, but vanishes everywhere except one point is a delta function[3], we can directly express the scalar product as

$$\langle p|p'\rangle = 2\pi\hbar\, \delta(p-p')$$

which proves Eq. 6.16.

Now it is a simple matter to determine the unity operator in momentum representation as given in Eq. 6.17. The form of the unity operator can be verified by showing that it does not change the state to which it is applied. It is sufficient to show this for an arbitrary state from a complete set of states, which we choose to be the momentum eigenstates $|p\rangle$.

Due to the superposition principle, it holds for any state if it holds for a complete set of states such as $|p\rangle$. The identity is shown as follows:

$$\hat{1}|p\rangle \stackrel{\text{Eq. 6.17}}{=} \int \frac{dp'}{2\pi\hbar}|p'\rangle\langle p'|p\rangle \stackrel{\text{Eq. 6.16}}{=} \int \frac{dp'}{2\pi\hbar}|p'\rangle 2\pi\hbar\delta(p-p') = |p\rangle$$

6.6 Application: Two-state system

After the harmonic oscillator, we learn here about the second most important model in quantum mechanics, the two-state system. The two-state system is useful as a model for any system where only the ground state and the first excited state is accessible. At low temperatures for example, it may happen that we only have to consider the lowest two states. There are also systems that only have two states. We will learn that the spin of an electron, which is the rotation of an electron about itself, can only have two states namely a clock-wise and a counter-clock wise rotation about an axis with a well defined absolute value of the angular momentum.

We denote the two states with the symbols $|\psi_0\rangle$ and $|\psi_1\rangle$. $|\psi_0\rangle$ is the ground state of the system and $|\psi_1\rangle$ is the first excited state.

Let us assume that they are eigenstates of the Hamilton operator with eigenvalues ϵ_0 for the ground state and ϵ_1 for the excited state. Therefore,

$$\begin{aligned}\hat{H}|\psi_0\rangle &= |\psi_0\rangle\epsilon_0 \\ \hat{H}|\psi_1\rangle &= |\psi_1\rangle\epsilon_1\end{aligned} \tag{6.20}$$

We will see later that we can always choose the eigenstates to be orthonormal, that is

$$\langle\psi_i|\psi_j\rangle = \delta_{i,j}$$

We can now express the Hamilton operator in the basis of its eigenstates, so that it is represented by a matrix

$$\begin{pmatrix}\langle\psi_0|\hat{H}|\psi_0\rangle & \langle\psi_0|\hat{H}|\psi_1\rangle \\ \langle\psi_1|\hat{H}|\psi_0\rangle & \langle\psi_1|\hat{H}|\psi_1\rangle\end{pmatrix} \stackrel{\text{Eq. 6.20}}{=} \begin{pmatrix}\epsilon_0 & 0 \\ 0 & \epsilon_1\end{pmatrix} =: \boldsymbol{H}$$

[3] This conclusion is not entirely strict because there are also distributions (functions) that are zero everywhere and integrate to zero. An example is the derivative of the delta function. An admixture of these functions cannot be excluded.

The Hamilton operator has the form

$$\hat{H} = \sum_{i,j=0}^{1} |\psi_i\rangle H_{i,j}\langle\psi_j| = |\psi_0\rangle\epsilon_0\langle\psi_0| + |\psi_1\rangle\epsilon_1\langle\psi_1| \tag{6.21}$$

Hence we can write the Hamilton matrix $\boldsymbol{H}$.

$$\mathbf{H} = \begin{pmatrix} \epsilon_0 & 0 \\ 0 & \epsilon_1 \end{pmatrix} = \frac{\epsilon_0+\epsilon_1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\epsilon_1-\epsilon_0}{2}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Without loss of generality, we can choose the energy zero to be equal with $\frac{\epsilon_0+\epsilon_1}{2}$, the mean value of the two energies, which deletes the first term in the above equation. Furthermore we introduce the frequency ω_0 via the excitation energy

$$\hbar\omega_0 \stackrel{\text{def}}{=} \epsilon_1 - \epsilon_0$$

so that

$$\boldsymbol{H} = \frac{1}{2}\hbar\omega_0\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Now we would like to create ladder operators just as for the harmonic oscillator they are defined by

$$\begin{aligned} \hat{a}|\psi_0\rangle &= |\varnothing\rangle \\ \hat{a}|\psi_1\rangle &= |\psi_0\rangle \end{aligned} \tag{6.22}$$

Note here that $|\varnothing\rangle \neq |\psi_0\rangle$! $|\varnothing\rangle$ is the zero-state and is associated with a wave function that is zero everywhere in space. $|\psi_0\rangle$ on the other hand is the ground state.

Let us evaluate a in the representation $(|\psi_0\rangle, |\psi_1\rangle)$

$$\begin{pmatrix} \langle\psi_0|\hat{a}|\psi_0\rangle & \langle\psi_0|\hat{a}|\psi_1\rangle \\ \langle\psi_1|\hat{a}|\psi_0\rangle & \langle\psi_1|\hat{a}|\psi_1\rangle \end{pmatrix} \stackrel{\text{Eq. 6.22}}{=} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

The states are orthonormal. We will later see that eigenstates of any hermitian operator (such as the Hamiltonian) to different eigenvalues are always orthogonal.

We can now evaluate the matrix elements of $a^\dagger$ as

$$\begin{pmatrix} \langle\psi_0|\hat{a}^\dagger|\psi_0\rangle & \langle\psi_0|\hat{a}^\dagger|\psi_1\rangle \\ \langle\psi_1|\hat{a}^\dagger|\psi_0\rangle & \langle\psi_1|\hat{a}^\dagger|\psi_1\rangle \end{pmatrix} = \begin{pmatrix} \langle\psi_0|\hat{a}|\psi_0\rangle^* & \langle\psi_1|\hat{a}|\psi_0\rangle^* \\ \langle\psi_0|\hat{a}|\psi_1\rangle^* & \langle\psi_1|\hat{a}|\psi_1\rangle^* \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Let us now investigate the commutator relations: We start working out the products of the ladder operators. We can use the matrix notation, if we stay in the representation of eigenstates.

$$\hat{a}^\dagger\hat{a} \;\hat{=} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad ; \quad \hat{a}\hat{a}^\dagger \;\hat{=} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\hat{a}^2 \;\hat{=} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \hat{=} \hat{\varnothing} \quad ; \quad \hat{a}^{\dagger 2} \;\hat{=} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \hat{\varnothing}$$

Hence we can write the Hamiltonian as

$$\hat{H} = \frac{1}{2}\hbar\omega_0[\hat{a}^\dagger, \hat{a}]_- \hat{=} \frac{1}{2}\hbar\omega_0\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

while the anti-commutator[4] is a constant

$$
\begin{aligned}
[\hat{a}, \hat{a}^\dagger]_+ &= \hat{1} \\
[\hat{a}, \hat{a}]_+ &= \hat{\emptyset} \\
[\hat{a}^\dagger, \hat{a}^\dagger]_+ &= \hat{\emptyset}
\end{aligned}
$$

Note here, that the commutator relations are similar to those of the harmonic oscillator as seen in Eq. 5.38 only that commutators and anti-commutators are interchanged. This will be important if we pass to a many particle description of Fermions such as the electron.

6.6.1 Pauli matrices

It can be shown that all matrices in two dimensional space can be expressed in terms of 4 Hermitian matrices, the three so-called **Pauli matrices** and the unity operator. Pauli matrices play an important role for the description of spin and in general for certain particles called Fermions. Therefore, we shall introduce them here. Some useful expressions that facilitate working with Pauli matrices are provided in App. R.

We can now look for the most general form of an observable in the space spanned by the two states. The most general Hermitian two-by-two matrix has the form

$$
A = \begin{pmatrix} a & b+ic \\ b-ic & d \end{pmatrix}
$$

which can be expressed in the form

$$
\begin{aligned}
A &= \frac{a+d}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - c\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \frac{a-d}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
&= \frac{a+d}{2} + b\sigma_x - c\sigma_y + \frac{a-d}{2}\sigma_z
\end{aligned}
$$

with real coefficients and the so-called **Pauli matrices** σ_x, σ_y and σ_z.

PAULI MATRICES

$$
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{6.23}
$$

The multiplication table, commutators and anti-commutators of Pauli matrices are derived in App. R. Since these relations will be useful for what follows I recommend to go through that appendix at this point.

While we started with observables, that is with hermitian matrices, it can be shown that any 2×2 matrix, and not only Hermitian matrices, can be expressed as superposition of the identity and the Pauli matrices with complex coefficients.

As an example, the ladder operators have the form

$$
\hat{a} = \frac{1}{2}(\hat{\sigma}_x - i\hat{\sigma}_y); \quad \hat{a}^\dagger = \frac{1}{2}(\hat{\sigma}_x + i\hat{\sigma}_y)
$$

and the Hamilton operator can be written as

$$
\hat{H} = \frac{1}{2}\hbar\omega\hat{\sigma}_z \tag{6.24}
$$

[4]See Eq. 5.34 on p. 96 for a definition of the anti-commutator

6.6.2 Excursion: The Fermionic harmonic oscillator (Home study)

This section is a fun chapter. There is no harm if you skip to the next section, at least for the moment. If you are still with me, lean back and enjoy! We will get a glimpse into the field of **super-symmetry**, which plays a role in elementary particle physics.

Super-symmetry is important because of the zero point motion. We will see later that when we quantize fields the zero point motion of the harmonic oscillator results in an infinite energy even for free space. At first this does not seem like a big problem, because we can always add a compensating energy. However, the theory of general relativity says that energy, which is a type of mass in this theory, distorts space and therefore results in gravitational forces. So, when we consider also gravitational forces, the zero-point energy is observable and we cannot simply get rid of it. One attempt to get rid of the zero-point energy is to introduce for every harmonic oscillator also a corresponding two-state system, so that their positive and negative zero point motions cancel out exactly. This is the basic idea of super-symmetry. We will later talk about particles, Bosons and Fermions, where the Fermions are particles that are represented by a two-state system in each point in space and Bosons are particles that are represented by harmonic oscillators. But now we have already gone way beyond what we learned sofar.

What we try to do here is to develop the two-state system analogously to the harmonic oscillator. In the two-state system and in the harmonic oscillator we used creation and annihilation operators to construct the eigenvalue spectrum. We have seen that a lot is similar but that commutators and anti-commutators interchange their role.

Here we follow up on that analogy between two-state systems and the harmonic oscillator and develop what is called the Fermionic harmonic oscillator. The main question we will ask is: What are the coordinate and momentum operators that describe the two-state system.

We build the momenta and position operators from odd and even combinations of the creation and annihilation operators roughly analogous to the ordinary harmonic oscillator.

$$\hat{\chi} = \sqrt{\frac{\hbar}{2}}(\hat{a}^\dagger + \hat{a})$$
$$\hat{\pi} = i\sqrt{\frac{\hbar}{2}}(\hat{a}^\dagger - \hat{a})$$

Both are Hermitian operators

Let us work out the product table

$$\hat{\chi}^2 = \frac{\hbar}{2} \;;\; \hat{\pi}^2 = \frac{\hbar}{2} \;;\; \hat{\pi}\hat{\chi} = +i\frac{\hbar}{2}\hat{\sigma}_z \;;\; \hat{\chi}\hat{\pi} = -i\frac{\hbar}{2}\hat{\sigma}_z \;;$$

in order to derive the commutation and anti-commutation relations

$[\hat{\chi}, \hat{\chi}]_+ = \hbar$	$[\hat{\chi}, \hat{\chi}]_- = 0$
$[\hat{\pi}, \hat{\pi}]_+ = \hbar$	$[\hat{\pi}, \hat{\pi}]_- = 0$
$[\hat{\pi}, \hat{\chi}]_+ = 0$	$[\hat{\pi}, \hat{\chi}]_- = i\hbar\sigma_z$

Thus we can express the Hamilton operator in the form

$$\hat{H} \stackrel{\text{Eq. 6.24}}{=} \frac{\hbar\omega}{2}\hat{\sigma}_z = \frac{-i\omega}{2}[\hat{\pi}, \hat{\chi}]_- \stackrel{[\hat{\pi},\hat{\chi}]_+=0}{=} i\omega\hat{\chi}\hat{\pi}$$

Now we see the analogy in full glory. It seems that all we need to do is to interchange plus and minus signs.

	harmonic oscillator	two-state system
(anti-)commutators	$[\hat{a}, \hat{a}^\dagger]_- = 1$	$[\hat{a}, \hat{a}^\dagger]_+ = 1$
	$[\hat{a}, \hat{a}]_- = [\hat{a}^\dagger, \hat{a}^\dagger]_- = 0$	$[\hat{a}, \hat{a}]_+ = [\hat{a}^\dagger, \hat{a}^\dagger]_+ = 0$
Hamilton operator	$\hat{H} = \frac{1}{2}\hbar\omega_0[\hat{a}^\dagger, \hat{a}]_+$	$\hat{H} = \frac{1}{2}\hbar\omega_0[\hat{a}^\dagger, \hat{a}]_-$
momentum	$\hat{p} = \frac{im\hbar\omega_0(\hat{a}^\dagger - \hat{a})}{\sqrt{2m\hbar\omega_0}}$	$\hat{\pi} = i\sqrt{\frac{\hbar}{2}}(\hat{a}^\dagger - \hat{a})$
position	$\hat{x} = \frac{\hbar(\hat{a}^\dagger + \hat{a})}{\sqrt{2m\hbar\omega_0}}$	$\hat{\chi} = \sqrt{\frac{\hbar}{2}}(\hat{a}^\dagger + \hat{a})$
annihilator	$\hat{a} = \frac{i\hat{p} + m\omega\hat{x}}{\sqrt{2m\hbar\omega_0}}$	$\hat{a} = \frac{i\hat{\pi} + \hat{\chi}}{\sqrt{2\hbar}}$
creator	$\hat{a}^\dagger = \frac{-i\hat{p} + m\omega\hat{x}}{\sqrt{2m\hbar\omega_0}}$	$\hat{a}^\dagger = \frac{-i\hat{\pi} + \hat{\chi}}{\sqrt{2\hbar}}$
(anti)-commutators	$[\hat{p}, \hat{x}]_- = \frac{\hbar}{i}$	$[\hat{\pi}, \hat{\chi}]_+ = 0$
	$[\hat{p}, \hat{p}]_- = [\hat{x}, \hat{x}]_- = 0$	$[\hat{\pi}, \hat{\pi}]_+ = [\hat{\chi}, \hat{\chi}]_+ = \hbar$
Hamilton operator	$\frac{\hat{p}^2}{2m} + m\omega^2\hat{x}^2$	$i\omega\hat{\chi}\hat{\pi}$

The classical limit is really funny. If we let $\hbar$ go to zero, we find that interchanging $\hat{\pi}$'s and χ's in a product still changes its sign. This is the so-called **Grassmann algebra**. In the Grassmann algebra we work with numbers, but any interchange of two numbers changes the sign of the product.

Why do we call this two-state system a harmonic oscillator? The regular harmonic oscillator has the property that the classical system has exactly one frequency. Let us therefore derive the classical Hamilton equations.

$$\dot{\chi} = \frac{\partial H}{\partial \pi} = -i\omega\chi \Rightarrow \chi(t) = \mathrm{e}^{-i\omega t}A$$
$$\dot{\pi} = -\frac{\partial H}{\partial \chi} = -i\omega\pi \Rightarrow \pi(t) = \mathrm{e}^{-i\omega t}B$$

(Note, that there seems to be a sign error in the first of the two last equation. This is not so, because we have to interchange χ and π in the Hamiltonian before we can form the derivative. Due to the anti-commutation relation for χ and π, this introduces a sign change.) Thus the "positions" χ and momenta perform oscillations with a single frequency just as the regular harmonic oscillator.

Chapter 7

Measurements

The main new feature of quantum mechanics as discussed so far was to replace particles by wave packets. Thus make the transition from a classical particle theory to a, still classical, field theory. The quantum mechanical measurement process is what is really puzzling about quantum mechanics. The rules are fairly clear, but there is still an ongoing discussion about its meaning.

In classical mechanics, measurable quantities are associated with functions $A(q,p)$ of the coordinates and momenta– or alternatively the velocities–in configuration space. A measurable quantity for a state is obtained by evaluating the value of the function A for this state. In quantum mechanics a measurable quantity is associated with an Hermitian operator.

Fig. 7.1: Niels Bohr, 1892-1987. Danish Physicist. Nobel price 1922 for his *investigations of the structure of atoms and the radiation emanating from them.*

7.1 Expectation values

Quantum theory does not predict the outcome of an experiment with certainty. However, it predicts the mean value of a measurable quantity when the experiment is repeated sufficiently often. This is less of a limitation than it seems, because we may very well measure the spread of the measured values. Hence, we can predict under which circumstances our prediction will be certain.

The mean value of a measured quantity is called **expectation value**. The expectation value of an operator A and a quantum mechanical state $|\psi\rangle$ is obtained as

$$\langle A\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} \tag{7.1}$$

In the following we assume that the state $|\psi\rangle$ is normalized, so that we can express the expectation value simply as

$$\langle A\rangle \stackrel{\langle\psi|\psi\rangle=1}{=} \langle\psi|\hat{A}|\psi\rangle$$

7.2 Certain measurements

If a state predicts a sharp value of an observable $\hat{A}$, the state is an eigenstate of $\hat{A}$.

Let us consider the **square variation**[1] σ^2 of an observable from its expectation value in some state. Given a state $|\psi\rangle$ we define the deviation $\Delta\hat{A}$ from the expectation value of that state.

$$\Delta\hat{A} \stackrel{\text{def}}{=} \hat{A} - \langle\psi|\hat{A}|\psi\rangle\hat{1} \tag{7.2}$$

Note that $\Delta\hat{A}$ describes one operator. The operator depends on the state $|\psi\rangle$. The expectation value in Eq. 7.2 is considered a fixed number. The square deviation is then the expectation value of $(\Delta\hat{A})^2$. Let us work out the expectation value of the square deviation

$$\begin{aligned}
\langle\psi|(\Delta\hat{A})^2|\psi\rangle &= \left\langle\psi|\left(\hat{A}-\langle\hat{A}\rangle\right)^2|\psi\right\rangle \\
&= \left\langle\psi|\left(\hat{A}-\frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}\right)^2|\psi\right\rangle \\
&= \langle\psi|\hat{A}^2|\psi\rangle - 2\langle\psi|\hat{A}|\psi\rangle\frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} + \langle\psi|\psi\rangle\left(\frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}\right)^2 \\
\frac{\langle\psi|(\Delta\hat{A})^2|\psi\rangle}{\langle\psi|\psi\rangle} &= \frac{\langle\psi|\hat{A}^2|\psi\rangle}{\langle\psi|\psi\rangle} - \left(\frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}\right)^2
\end{aligned} \tag{7.3}$$

of a measurable quantity in an experiment for a given wave function.

Using the expression Eq. 7.1 for the expectation value we find

$$\langle(\Delta\hat{A})^2\rangle \stackrel{\text{Eq. 7.3}}{=} \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2$$

This result reminds strongly of the classical expression for the square deviation σ_A^2 of a quantity A. We consider here an experiment, which contains some randomness. The measurement can produce any of the possible values a_i of A. The frequency of the outcome a_i is P_i, which is the number of times a_i was the measured result relative to the total number of measurements. The frequency P_i can be interpreted as probability of the outcome a_i. The mean value of the measured quantity A is

$$\langle A\rangle \stackrel{\text{def}}{=} \sum_i P_i a_i$$

and the square deviation is

$$\begin{aligned}
\sigma_A^2 = \langle(A-\langle A\rangle)^2\rangle &= \sum_i P_i(a_i - \langle A\rangle)^2 = \sum_i P_i(a_i^2 - 2a_i\langle A\rangle + \langle A\rangle^2) \\
&= \underbrace{\sum_i P_i a_i^2}_{\langle A^2\rangle} - 2\langle A\rangle \underbrace{\sum_i P_i a_i}_{\langle A\rangle} + \langle A\rangle^2) \\
&= \langle A^2\rangle - \langle A\rangle^2
\end{aligned}$$

After this little detour we return back to quantum mechanics.

Since the operator is Hermitian – it corresponds to a measurable quantity –, the condition for a certain prediction without spread is

$$\langle(\Delta\hat{A})^2\rangle = 0$$

which implies

$$\begin{aligned}
\langle(\Delta\hat{A})^2\rangle &\stackrel{\text{Eq. 7.3}}{=} \langle\psi|\hat{A}^2\psi\rangle\langle\psi|\psi\rangle - \langle\hat{A}\psi|\psi\rangle^2 \\
\Rightarrow \quad \underbrace{\langle\hat{A}\psi|\hat{A}\psi\rangle}_{\langle\phi|\phi\rangle}\langle\psi|\psi\rangle &= \underbrace{\langle\hat{A}\psi|\psi\rangle}_{\langle\phi|\psi\rangle}\underbrace{\langle\psi|\hat{A}\psi\rangle}_{\langle\psi|\phi\rangle}
\end{aligned}$$

[1]germ: Quadratische Abweichung

Let us introduce $|\phi\rangle \stackrel{\text{def}}{=} \hat{A}|\psi\rangle$. **Schwarz' inequality**, which is derived below in 7.4, says that for any two states $|\phi\rangle$ and $|\psi\rangle$ the inequality

$$\langle\phi|\phi\rangle\langle\psi|\psi\rangle - \langle\phi|\psi\rangle\langle\psi|\phi\rangle \stackrel{\text{Eq. 7.4}}{\geq} 0$$

holds and that the equal sign holds, if the two states $|\psi\rangle$ and $|\phi\rangle$ are linear dependent, i.e. $|\phi\rangle = |\psi\rangle a$, with some proportionality factor a.

For our problem with $\langle(\Delta\hat{A})^2\rangle = 0$, Schwarz' inequality implies

$$\hat{A}|\psi\rangle = |\psi\rangle a$$

We call such an equation an eigenvalue equation. The state $|\psi\rangle$ is called **eigenstate**, and the value a is called an **eigenvalue**. Thus we have shown that a sharp measurement is only possible for an eigenstate of the corresponding operator.

7.2.1 Schwarz' inequality

Let us now derive **Schwarz' inequality** we used before.

SCHWARZ' INEQUALITY

Schwarz' inequality says that for two arbitrary states $|\phi\rangle$ and $|\psi\rangle$

$$\langle\phi|\phi\rangle\langle\psi|\psi\rangle - \langle\phi|\psi\rangle\langle\psi|\phi\rangle \geq 0 \qquad (7.4)$$

and that the equality holds only when $|\phi\rangle$ and $|\psi\rangle$ are linearly dependent, i.e. if $|\phi\rangle = |\psi\rangle a$ with some (complex) proportionality constant a.

The essence of Schwarz inequality is that a vector becomes shorter if we project something out of it. The above relation can be written as

$$\langle\psi|\left(\hat{1} - |\phi\rangle\frac{1}{\langle\phi|\phi\rangle}\langle\phi|\right)|\psi\rangle \stackrel{\text{Eq. 7.4}}{\geq} 0$$

In order to grasp the meaning of Schwarz' inequality consider regular vectors instead of quantum mechanical states. For regular vectors Schwarz inequality says

$$\vec{a}^2\vec{b}^2 - (\vec{a}\vec{b})^2 \geq 0$$

Dividing this expression by the squared length of the two vectors yields

$$1 - \left(\frac{\vec{a}\,\vec{b}}{|\vec{a}||\vec{b}|}\right)^2 = 1 - \cos^2(\alpha) = \sin^2(\alpha) \geq 0$$

where α is the angle between the two vectors. Clearly, the squared sinus is always positive and it vanishes for an angle of 0° and 180°, that is for parallel or antiparallel vectors. The analogy with regular vectors may be instructive in the proof of Schwarz' inequality for quantum mechanical states.

The proof of Schwarz inequality goes as follows:

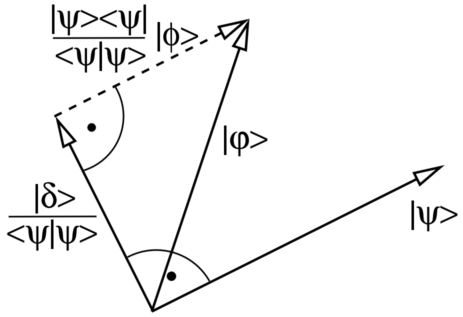
PROOF of Schwarz inequality Eq. 7.4: Let us say we have two states $|\psi\rangle$ and $|\phi\rangle$. Out of the two states, we construct one, $|\delta\rangle = |\phi\rangle\langle\psi|\psi\rangle - |\psi\rangle\langle\psi|\phi\rangle$, that is orthogonal to $|\psi\rangle$. We express $|\phi\rangle$ as

$$|\phi\rangle = \Big[|\delta\rangle + |\psi\rangle\langle\psi|\phi\rangle\Big]\frac{1}{\langle\psi|\psi\rangle}$$

and evaluate its norm exploiting the orthogonality $\langle\psi|\delta\rangle = 0$

$$\langle\phi|\phi\rangle = \Big[\langle\delta|\delta\rangle + \langle\phi|\psi\rangle\langle\psi|\psi\rangle\langle\psi|\phi\rangle\Big]\frac{1}{\langle\psi|\psi\rangle^2}$$

$$\Rightarrow \langle\phi|\phi\rangle\langle\psi|\psi\rangle - \langle\phi|\psi\rangle\langle\psi|\phi\rangle = \frac{\langle\delta|\delta\rangle}{\langle\psi|\psi\rangle}$$



1. Since the norm of every state is positive or zero, the right hand side is positive or zero.
2. The right hand side vanishes, when $|\delta\rangle = 0$. In that case, however, we find from the equation defining $|\delta\rangle$ that

$$|\phi\rangle = |\psi\rangle\frac{\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle}$$

   Hence $|\phi\rangle$ and $|\psi\rangle$ are linearly dependent if the equal sign holds.

q.e.d.

7.3 Eigenstates

We have learned that the uncertainty of a measurement vanishes, if the state is an eigenstate of the operator corresponding to the measured quantity. Eigenstates play an important role in measurements.

There are a few important points to know about eigenstates:

1. The eigenvalues of a Hermitian operator are real.

PROOF: Let ψ be an eigenstate of the Hermitian operator $\hat{A}$ with eigenvalue a.

$$\langle\psi|\psi\rangle a = \langle\psi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}^\dagger|\psi\rangle^* \stackrel{\hat{A}=\hat{A}^\dagger}{=} \langle\psi|\hat{A}|\psi\rangle^* = \langle\psi|\psi\rangle^* a^*$$

Because the norm of a state is real, we find that the eigenvalue $a = a^*$ is real, too.
q.e.d

2. The eigenstates of a Hermitian operator with different eigenvalues are orthogonal.

PROOF: Let us consider two eigenstates $|\psi_1\rangle$ and $\psi_2\rangle$ of a Hermitian operator $\hat{A}$ with different eigenvalues a_1 and a_2.
Because $\hat{A}$ is Hermitian we obtain

$$0 = \langle\phi|\hat{A} - \hat{A}^\dagger|\psi\rangle$$

We continue by using the eigenvalue equations $\hat{A}|\psi_1\rangle = |\psi_1\rangle a_1$ and $\hat{A}|\psi_2\rangle = |\psi_2\rangle a_2$.

$$\begin{aligned}
0 &= \langle\psi_1|\hat{A} - \hat{A}^\dagger|\psi_2\rangle \\
&= \langle\psi_1|\hat{A}|\psi_2\rangle - \langle\psi_2|\hat{A}|\psi_1\rangle^* \\
&= \langle\psi_1|\psi_2\rangle a_2 - \langle\psi_2|\psi_1\rangle^* a_1^* \\
&= \langle\psi_1|\psi_2\rangle[a_2 - a_1]
\end{aligned}$$

(Note that $a_1^* = a_1$, since $\hat{A}$ is Hermitian.) If the eigenvalues are different this equation can only be fulfilled if the scalar product vanishes, thus the two eigenstates are orthogonal.
q.e.d.

3. Any linear combination of eigenstates of an operator with the same eigenvalue is itself an eigenstate with this eigenvalue.

PROOF: We start with two eigenstates to $\hat{A}$ with the same eigenvalue a

$$\begin{aligned}
\hat{A}[|\psi_1\rangle c_1 + |\psi_2\rangle c_2] &= [\hat{A}|\psi_1\rangle c_1 + \hat{A}|\psi_2\rangle c_2] = [|\psi_1\rangle ac_1 + |\psi_2\rangle ac_2] = \\
&= [|\psi_1\rangle c_1 + |\psi_2\rangle c_2]a
\end{aligned}$$

q.e.d.

4. The eigenstates to the same eigenvalue can be orthogonalized. The number of orthogonal eigenstates for a given eigenvalue is called the **degeneracy** of that eigenvalue.

PROOF: This is a trivial statement, because any set of states can be orthogonalized. Here we show a procedure to do this in practice. The method is called **Gram-Schmidt orthogonalization**.
We start out from a set of states $\{|\psi_i\rangle\}$ and orthogonalize them recursively. The resulting states $\{|\phi_i\rangle\}$ will not only be orthogonalized but also normalized, so that $\langle\phi_i|\phi_j\rangle = \delta_{i,j}$.

$$\begin{aligned}
|\xi_n\rangle &= |\psi_n\rangle - \sum_{j=1}^{n-1} |\phi_j\rangle\langle\phi_j|\psi_n\rangle \\
|\phi_n\rangle &= |\xi_n\rangle \frac{1}{\sqrt{\xi_n|\xi_n\rangle}}
\end{aligned}$$

In a first step, the orthogonalization step, we produce the states $|\xi_n\rangle$. Only those states $|\phi_n\rangle$ are required that have already been constructed. We can assume that the states $|\phi_j\rangle$ are normalized, because, in the second step, the normalization step, the orthonormalized state $|\phi\rangle$ is obtained by scaling $|\xi_n\rangle$. The recursive equation can produce a zero state $|\xi_n\rangle = 0$, which produces a divide-by-zero in the next step. When this happens $|\psi_n\rangle$ is linearly dependent to the other states and must be excluded from the set. Thus, the orthogonal set $\{|\phi_j\rangle\}$ may have less elements than the initial set $\{|\psi_i\rangle\}$. *q.e.d.*

The set of all eigenvalues of an operator $\hat{A}$ is called the **spectrum** of $\hat{A}$. The spectrum can be partly discrete and partly continuous. We denote the eigenstates with discrete eigenvalues with in integer index $|\phi_i\rangle$ and those from the continuous spectrum $|\phi_\alpha\rangle$ with a continuous index α.

We normalize the states that

$$\langle\phi_i|\phi_j\rangle = \delta_{i,j}$$
$$\langle\phi_\alpha|\phi_{\alpha'}\rangle = \delta(\alpha-\alpha')$$
$$\langle\phi_i|\phi_\alpha\rangle = 0$$

If the orthogonal system of eigenstates is complete, that is every state $|\psi\rangle$ can be represented as superposition of eigenstates ϕ_i,

$$|\psi\rangle = \sum_i |\phi_i\rangle c_i + \int d\alpha |\phi_\alpha\rangle c_\alpha$$

then the Hermitian operator is called an **observable**. The coefficients are $c_i = \langle\phi_i|\psi\rangle$ and $c_\alpha = \langle\phi_\alpha|\psi\rangle$.

Important Note! In the following we will often use the sum alone even if the eigenvalue spectrum may be completely or partly continuous. In this case, we always imply summation over all states of the discrete part of the spectrum and integration over all states from the continuous spectrum.

We have already made contact with operators with discrete and with continuous eigenvalue spectra. The Hamiltonian of the harmonic oscillator has a purely discrete eigenvalue spectrum. The Hamiltonian for the free particle on the other hand has a purely continuous eigenvalue spectrum.

7.4 Heisenberg's uncertainty relation

HEISENBERG'S UNCERTAINTY RELATION

Heisenberg's uncertainty principle [27] says that two canonical conjugate quantities such as position and momentum cannot be determined simultaneously with absolute accuracy. Furthermore it gives a lower limit

$$\sqrt{\langle\psi|(\Delta\hat{Q})^2|\psi\rangle}\sqrt{\langle\psi|(\Delta\hat{P})^2|\psi\rangle} \geq \frac{\hbar}{2} \qquad (7.5)$$

for the certainty with which both quantities can be determined.
Here we used the operator $\Delta\hat{A}$ that describes the deviation of an operator $\hat{A}$ from its mean value $\langle\psi|\hat{A}|\psi\rangle$ is

$$\Delta\hat{A} \stackrel{\text{def}}{=} \hat{A} - \langle\psi|\hat{A}|\psi\rangle \qquad (7.6)$$

The uncertainty principle can be generalized[28] to a general pair of observables $\hat{A}$ and $\hat{B}$ saying

$$\sqrt{\langle\psi|(\Delta\hat{A})^2|\psi\rangle}\sqrt{\langle\psi|(\Delta\hat{B})^2|\psi\rangle} \geq \frac{1}{2}\Big|\langle\psi|\ [A,B]\ |\psi\rangle\Big| \qquad (7.7)$$

Two quantities are canonical conjugate of each other if they are both Hermitian and if their commutator is $[\hat{P},\hat{Q}] = \frac{\hbar}{i}$.

As shown earlier, we define the mean square deviation of an observable from the mean value of an operator $\hat{Q}$ in a state $|\psi\rangle$ as

$$\sqrt{\langle\psi|(\Delta\hat{Q})^2|\psi\rangle} \stackrel{\text{Eq. 7.6}}{=} \sqrt{\langle\psi|(\hat{Q}-\langle\psi|\hat{Q}|\psi\rangle)^2|\psi\rangle} = \sqrt{\langle\psi|\hat{Q}^2|\psi\rangle - \langle\psi|\hat{Q}|\psi\rangle^2}$$

Heisenberg's uncertainty relation says that one cannot prepare a system in a state such that two observables have sharp values, unless their operator vanishes.

The relevance of Heisenberg's uncertainty principle reaches further than only quantum mechanics. It is important for optics and for signal processing in general. Translated into optics it says that we can only make a light-pulse shorter at the cost of its frequency resolution. A femtosecond light pulse will have a wide frequency distribution. In signal processing it says that, if our frequency band for transmission is limited, this also limits the sharpness of the transmitted signal in the time domain.

The underlying principle is that the Fourier transform of a wave packet, becomes wider as we decrease the width of the wave packet. This is an intrinsic property of a Fourier transform. Two conjugate variables can always be considered as the variables in a Fourier transform. For example

$$\underbrace{\langle x|\psi\rangle}_{\psi(x)}=\langle x|\underbrace{\int\frac{dp}{2\pi\hbar}\,|p\rangle\langle p|}_{\hat{1}}\psi\rangle=\int\frac{dp}{2\pi\hbar}\,\underbrace{\langle x|p\rangle}_{\mathrm{e}^{\frac{i}{\hbar}px}}\langle p|\psi\rangle=\int\frac{dp}{2\pi\hbar}\,\mathrm{e}^{\frac{i}{\hbar}px}\,\underbrace{\langle p|\psi\rangle}_{\psi(p)}$$

Since time and frequency ω are also related by a Fourier transform, just like position and momenta, it is clear that the results of Heisenberg's uncertainty principle can be translated into the above problems in the time-frequency domain.

PROOF: The proof of Heisenberg's uncertainty principle goes as follows: For two observable $\hat{A}$ and $\hat{B}$ let us define the corresponding deviation as

$$\Delta\hat{A}\stackrel{\text{def}}{=}\hat{A}-\langle\hat{A}\rangle\qquad\text{and}\qquad\Delta\hat{B}\stackrel{\text{def}}{=}\hat{B}-\langle\hat{B}\rangle$$

Next, we introduce two states, namely

$$|\phi_A\rangle\stackrel{\text{def}}{=}\Delta\hat{A}|\psi\rangle\qquad\text{and}\qquad|\phi_B\rangle\stackrel{\text{def}}{=}\Delta\hat{B}|\psi\rangle$$

We express the product of the deviations by the states $|\phi_A\rangle$ and $|\phi_B\rangle$. Then we exploit Schwarz' inequality Eq. 7.4 to obtain a lower bound.

$$\begin{aligned}\langle\psi|(\Delta\hat{A})^2|\psi\rangle\langle\psi|(\Delta\hat{B})^2|\psi\rangle&\stackrel{\Delta\hat{A}=\Delta\hat{A}^\dagger}{=}\underbrace{\langle\psi|(\Delta\hat{A})^\dagger}_{\langle\phi_A|}\underbrace{(\Delta\hat{A})|\psi\rangle}_{|\phi_A\rangle}\underbrace{\langle\psi|(\Delta\hat{B})^\dagger}_{\langle\phi_B|}\underbrace{(\Delta\hat{B})|\psi\rangle}_{|\phi_B\rangle}\\&=\langle\phi_A|\phi_A\rangle\langle\phi_B|\phi_B\rangle\stackrel{\text{Eq. 7.4}}{\geq}|\langle\phi_A|\phi_B\rangle|^2\\&=\Big|\underbrace{\langle\psi|(\Delta\hat{A})^\dagger}_{\langle\phi_A|}\underbrace{(\Delta\hat{B})|\psi\rangle}_{|\phi_B\rangle}\Big|^2\\&=\Big|\langle\psi|\frac{1}{2}[\Delta\hat{A},\Delta\hat{B}]_-+\frac{1}{2}[\Delta\hat{A},\Delta\hat{B}]_+|\psi\rangle\Big|^2\end{aligned}\tag{7.8}$$

We wrote the product in terms of commutator and anti-commutator, because one can show that the expectation value of the commutator is purely imaginary, and that of the anticommutator is purely real. This fact will be useful in the following and therefore we derive it here: We begin with

$$\langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle=\langle\psi|\Big(\Delta\hat{B}^\dagger\Delta\hat{A}^\dagger\Big)^\dagger|\psi\rangle=\langle\psi|\Delta\hat{B}^\dagger\Delta\hat{A}^\dagger|\psi\rangle^*=\langle\psi|\Delta\hat{B}\Delta\hat{A}|\psi\rangle^*$$

where we exploited only that the two operators are hermitian. Now, we can show that the expectation values of commutator and anticommutator are either purely real or purely imaginary.

$$\begin{aligned}\langle\psi|[\Delta\hat{A},\Delta\hat{B}]_-|\psi\rangle&=\langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle-\langle\psi|\Delta\hat{B}\Delta\hat{A}|\psi\rangle=\langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle-\langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle^*\\&=2i\mathrm{Im}\Big(\langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle\Big)\end{aligned}\tag{7.9}$$

$$\begin{aligned}\langle\psi|[\Delta\hat{A},\Delta\hat{B}]_+|\psi\rangle&=\langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle+\langle\psi|\Delta\hat{B}\Delta\hat{A}|\psi\rangle=\langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle+\langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle^*\\&=2\mathrm{Re}\Big(\langle\psi|\Delta\hat{A}\Delta\hat{B}|\psi\rangle\Big)\end{aligned}\tag{7.10}$$

...cont'd

cont'd...
With Eqs. 7.9,7.10 we can express the absolute square by the squares of real and imaginary part.

$$
\begin{aligned}
\langle\psi|(\Delta\hat{A})^2|\psi\rangle\langle\psi|(\Delta\hat{B})^2|\psi\rangle &\stackrel{\text{Eq. 7.8}}{=} \left|\langle\psi|\frac{1}{2}[\Delta\hat{A},\Delta\hat{B}]_-|\psi\rangle + \langle\psi|\frac{1}{2}[\Delta\hat{A},\Delta\hat{B}]_+|\psi\rangle\right|^2 \\
&\stackrel{\textit{Eqs. } 7.9,7.10}{=} \left|\langle\psi|\frac{1}{2}[\Delta\hat{A},\Delta\hat{B}]_-|\psi\rangle\right|^2 + \left|\langle\psi|\frac{1}{2}[\Delta\hat{A},\Delta\hat{B}]_+|\psi\rangle\right|^2 \\
&= \left|\langle\psi|\frac{1}{2}[\hat{A},\hat{B}]_-|\psi\rangle\right|^2 + \left|\langle\psi|\frac{1}{2}[\Delta\hat{A},\Delta\hat{B}]_+|\psi\rangle\right|^2 \qquad (7.11)
\end{aligned}
$$

In the last expression we exploited that the commutator of the deviations is the same as that of the original operators. A similar simplification is not possible for the anticommutator.
The extended form of the uncertainty principle from Eq. 7.11 is due to Schrödinger[29]. It is probably the most stringent form of the uncertainty principle.
By dropping the term with the anticommutator, we obtain the even lower bound specified in Eq. 7.7.

$$\sqrt{\langle\psi|(\Delta\hat{A})^2|\psi\rangle}\sqrt{\langle\psi|(\Delta\hat{B})^2|\psi\rangle} \geq \frac{1}{2}\left|\langle\psi|[\hat{A},\hat{B}]_-|\psi\rangle\right|$$

and by insertion of the commutator of canonical conjugate variables $[P,Q]_- = \frac{\hbar}{i}$ one obtains the special form of Heisenberg Eq. 7.5.
q.e.d

7.5 Measurement process

Now we are prepared to investigate the measurement process. The measurement process in quantum mechanics exhibits some surprising elements. I will approach this topic as follows: Firstly, I will describe the so-called Kopenhagen interpretation[2] of the measurement process, which goes back to Niels Bohr. The resulting rules can easily be memorized and they capture everything needed for the everyday use of quantum mechanics.

The Kopenhagen interpretation had some ad-hoc definitions that were not quite satisfactory. One of these problems leads to puzzles such as the one named **Schrödinger's cat**. The **theory of decoherence** digs deeper into the measurement process and arrives at a conceptionally satisfactory description of the measurement process.

7.6 Kopenhagen interpretation

The quantum-mechanical measurement process can be described by postulates, which are part of the conceptual basis of quantum mechanics.

If we perform a measurement of an observable $\hat{A}$, that is denoted by a Hermitian operator, the outcome will be value a. After the measurement, we know the value of a with certainty. Hence, the state must be in an eigenstate of the operator $\hat{A}$ with eigenvalue a. This tells us a few important lessons, which are postulates of quantum mechanics:

- **Spectrum of measurable values:** The only values, one can possibly obtain by measuring a quantity related to an operator $\hat{A}$, are its eigenvalues. If the spectrum of $\hat{A}$ is discrete, one can only obtain certain discrete values.

- **Probabilistic interpretation of quantum mechanics:** Quantum mechanics does not predict which one of the eigenvalues of $\hat{A}$ is obtained in a particular measurement, but it predicts the

[2]See Copenhagen Interpretation of Quantum Mechanics (Stanford Encyclopedia of Philosophy) https://plato.stanford.edu/entries/qm-copenhagen/.

probability with which one obtains a certain value by $P_a = \sum_i \langle\Psi|\chi_i\rangle\langle\chi_i|\Psi\rangle$, where the sum runs again over all eigenstates $|\chi_i\rangle$ with eigenvalue a.

- **Collapse of the wave function:** During the measurement, the wave function "collapses". The measurement process changes the wave function into an eigenstate of $\hat{A}$. If the wave function before the measurement was $|\Psi\rangle$, after the measurement it is $\frac{\sum_i |\chi_i\rangle\langle\chi_i|\Psi\rangle}{\sqrt{\sum_i |\langle\chi_i|\Psi\rangle|^2}}$, where $|\chi_i\rangle$ are normalized eigenstates of $\hat{A}$ to the eigenvalue of a.

Let us describe the process of performing a laboratory experiment including a measurement in more detail:

1. Set up the experimental conditions, that is define the Hamilton operator $\hat{H}(t)$
2. Define the initial conditions. A certain state $|\psi(t_0)\rangle$ is prepared
3. The system is allowed to evolve. The dynamics is governed by Schrödinger's equation. At some point in time, t_1 the actual measurement takes place. Just before the measurement, the state is $|\psi(t_1^-)\rangle$. The minus sign indicates that we consider a time arbitrary close to that of the measurement, but still before the actual measurement takes place.
4. The measurement is performed. We obtain one of the possible outcomes a_i. Even though we could conceive a larger variety of outcomes, quantum mechanics says that we will obtain only those numbers, that correspond to eigenvalues of the measured quantity $\hat{A}$. If we repeat the experiment, we do not obtain the same value for A each time the experiment is performed, but we obtain the eigenvalues of $\hat{A}$ in a seemingly random manner. Thus, the experiment is not deterministic. While we cannot describe the experiment by a single outcome, we can determine the frequencies P_i of outcomes a_i. The **frequency** is the ratio of experiments with a particular outcome a_i. Quantum theory predicts the frequencies as

$$P_i = |\langle\chi_i|\psi\rangle|^2$$

   where $|\chi_i\rangle$ is a normalized eigenstate of $\hat{A}$ with eigenvalue a_i, that is $\hat{A}|\chi_i\rangle = |\chi_i\rangle a_i$. If the eigenvalue is degenerate, the frequency is obtained as the sum over all states with that eigenvalue.
5. After each measurement, we can perform additional measurements of the same or different observable. The system behaves as if the state would be changed during the measurement abruptly from $|\psi(t_1^-)\rangle$ to

$$|\psi(t_1^+)\rangle = |\chi_i\rangle \frac{\langle\chi_i|\psi(t_1^-)\rangle}{\sqrt{|\langle\chi_i|\psi(t_1^-)\rangle|^2}}$$

   The term in the denominator ensures that the new state is again normalized. Thus the state $|\psi(t_1^-)\rangle$ is projected onto the eigenstate of $\hat{A}$ that was the outcome of the experiment.

   If the eigenvalue a_i is degenerate we need to include all projections onto eigenstates with the same eigenvalue, i.e.

$$|\psi(t_1^+)\rangle = \frac{\sum_{j:\hat{A}|\chi_j\rangle=|\chi_j\rangle a_i} |\chi_j\rangle\langle\chi_j|\psi(t_1^-)\rangle}{\sqrt{\sum_{j:\hat{A}|\chi_j\rangle=|\chi_j\rangle a_i} |\langle\chi_j|\psi(t_1^-)\rangle|^2}}$$

Repeated measurements

If the measurement of $\hat{A}$ is repeated twice on the same system, we obtain the same result, namely a, with certainty . Something unexpected happens, if we measure two different observables $\hat{A}$ and $\hat{B}$

one after the other. Unless the two operators commutate, implying that there is a set of common eigenstates, every measurement of the variable $\hat{B}$ destroys the information of $\hat{A}$ and vice versa. If we measure $\hat{A}$ twice on the same system, the result depends if we measured a quantity $\hat{B}$ between the two measurements of $\hat{A}$ or not.

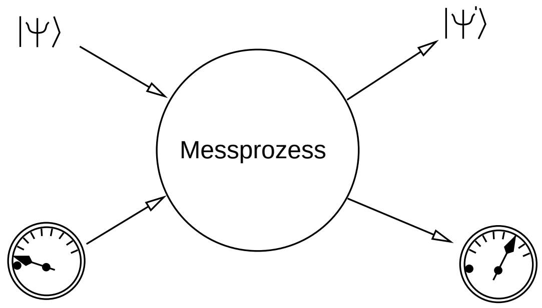
7.7 Decoherence

The Kopenhagen interpretation is a completely satisfactory working solution to the measurement process. The predictions are undisputed. Nevertheless, it has one unsatisfactory aspect: While the dynamics is described by the Schrödinger equation, a measurement causes an irreversible collapse of the wave function that is in direct conflict with the Schrödinger equation.

Decoherence theory investigates the measurement process in more detail and describes how seemingly classical behavior results from the dynamics of the measurement process in full accord with the Schrödinger equation. The "trick" is to consider our system as an open system, that is in contact with an environment. This interaction has to fulfill certain criteria in order to introduce irreversibility, such that a description with classical alternatives for the outcome of the measurement results. An excellent introduction to decoherence theory is given by Zurek[30].

Decoherence divides the measurement into several parts, each of which can be attributed to a well-defined physical process. In the following, I will not present a formal derivation, but only summarize my view on it.

- Let us start out with an initial wave function $|\psi_1\rangle = \sum_\sigma |\sigma\rangle\alpha_\sigma$ for our system, which may be in a superposition of eigenstates of the observable $\hat{A}$ that we wish to measure, i.e. $\hat{A}|\sigma\rangle = |\sigma\rangle\sigma$.



- **Transfer of the information:** We bring the system described by the wave function $|\psi\rangle$ in contact with a detector d, which can be in any of the states $|d_\sigma\rangle$. The interaction must be exactly such that, after the measurement, the combined system of measured system and the detector is

$$|\psi_2\rangle = \sum_\sigma |\sigma, d_\sigma\rangle\alpha_\sigma$$

  Now the detector contains all the information of the measured system, but it is itself in a superposition of states. One says that the detector is **entangled** with the system to be measured. An important concept, often missed, is that a superposition of states is different from a set of states with some probabilities, i.e. an **ensemble of states**. A superposition of states is still a single wave function. A set of states describes different alternatives for the wave function.

- **Decoherence:** An integral condition for a detector is that it is coupled to an environment. There must be a small but non-trivial interaction of the detector with the environment. If we then investigate our local system, it behaves just like an ensemble of states with probabilities equal to $P_\sigma = |\alpha_\sigma|^2$. This is an important step, because, from now on, our description is one that gives each eigenvalue of the observable a classical probability.

- **Collapse:** From now on, the situation is analogous to what happens with a classical system for which we do not have the full information. We describe the system with a probability distribution. By performing the measurement we gain additional information, so that we choose a better description by replacing the possible outcomes of the measurement by only one of the alternatives, namely the one that has been measured. This is the collapse of the probability distribution.

7.8 Difference between a superposition and a statistical mixture of states

The essence of the measurement process is that a superposition of states is converted into a weighted mixture of states. In the Kopenhagen interpretation this is done by a postulate, while in decoherence theory this is due to a physical process that results from a specific coupling of the measurement apparatus to an environment.

The distinction between a superposition of states and an mixture of states is however a bit subtle, which is the reason for discussing it in greater detail.

7.8.1 Mixture of states

Let us first explain a statistical mixture: Imagine that we do not know the state of the system. In order to describe the situation, we may provide several guesses for the state and attribute a probability to each of them. We may also select a complete set of eigenstates $|a_i\rangle$ of an observable $\hat{A}$ and provide each with a probability.

In order to obtain the expectation value of the observable $\hat{A}$ we can weight the expectation value of each state with the probability of the state, i.e.

$$\langle A\rangle=\sum_i P_i\langle a_i|\hat{A}|a_i\rangle=\sum_i P_i a_i$$

We obtain the weighted average of the eigenvalues of $\hat{A}$

However, we have accidentally considered the expectation value of a special observable, namely the one which provided the states to describe our mixture. For another observable $\hat{B}$, we obtain

$$\langle B\rangle=\sum_i P_i\langle a_i|\hat{B}|a_i\rangle$$

The result looks similar. Instead of the eigenvalues of B we have to use the expectation values of $\hat{B}$ with the set of states from our mixture. The probabilities are the same as those that we used to determine the expectation value of the observable $\hat{A}$.

7.8.2 Superposition of states

Let us now compare these equations for the expectation value with those from a superposition of states.

We start with a quantum-mechanical state $|\Psi\rangle$. If we want to measure a certain observable $\hat{A}$ it is convenient to expand this state into (orthonormal) eigenstates $|a_i\rangle$ of the operator $\hat{A}$, i.e.

$$|\psi\rangle=\underbrace{\sum_i|a_i\rangle\langle a_i|}_{\hat{1}}\psi\rangle=\sum_i|a_i\rangle\underbrace{c_i}_{\langle a_i|\psi\rangle}$$

This is a superposition of states $|a_i\rangle$. The components of the wave function are **coherent**.

The expectation value of the observable $\hat{A}$ for such a superposition is

$$\langle\psi|\hat{A}|\psi\rangle = \sum_{i,j} c_i^* c_j \langle a_i|\hat{A}|a_j\rangle = \sum_{i,j} c_i^* c_j \langle a_i|a_j\rangle a_j = \sum_i c_i^* c_i a_i = \sum_i P_i^{(A)} a_i$$

Thus the expectation value can be expressed as a sum over the eigenvalues of A with probabilities $P_i = c_i^\dagger c_i = \langle\psi|a_i\rangle\langle a_i|\psi\rangle$ for each eigenvalue.

This equation is basically identical to the one obtained for a mixture of states. However, let us now look at the expectation value of some other operator $\hat{B}$.

$$\langle\psi|\hat{B}|\psi\rangle = \sum_{i,j} c_i^* c_j \langle a_i|\hat{B}|a_j\rangle = \sum_i c_i^* c_i \langle a_i|\hat{B}|a_i\rangle + \sum_{i\neq j} c_i^* c_j \langle a_i|\hat{B}|a_j\rangle$$

This expression differs from the expression for the mixture by the presence of the off-diagonal terms of the double sum. These off-diagonal terms are a sign of a **coherent** superposition of states. Coherent means the following: For a given set of probabilities $P_i^{(A)}$, we can determine the coefficients

$$c_i = \sqrt{P}_i e^{i\varphi_i}$$

only if we also specify the relative phases φ_i of the states.

While a mixture of states is uniquely defined by a set of states, say $\{|a_i\rangle\}$ and their probabilities P_i, a superposition carries additional information, namely the phase information.

If we average the expectation value for $\hat{B}$ over all states that only differ by the relative phases, the off-diagonal terms cancel out, and we are left with an expression for expectation values that is identical to that for a mixture. This is the essence of the so-called **random-phase approximation**. The random phase approximation converts a superposition of states into a mixture of states.

It may be instructive to consider the expectation value for a superposition of states in terms of eigenvalues of the observable B. Let us simply introduce the eigenstates $|b_i\rangle$ and the corresponding eigenvalues b_i.

$$\langle\psi|\hat{B}|\psi\rangle = \sum_{i,j} \langle\psi|b_i\rangle\langle b_i|\hat{B}|b_j\rangle\langle b_j|\psi\rangle = \sum_i |\langle\psi|b_i\rangle|^2 b_i = \sum_i P_i^{(B)} b_i$$

While we can express the expectation value of B as a weighted sum of eigenvalues, the probabilities $P_i^{(B)}$ differ from the probabilities $P_i^{(A)}$, we have used before to obtain the expectation value of A.

7.8.3 Implications for the measurement process

We have seen that there is a fundamental difference between a mixture and a superposition of states. An instrument can only be considered a detector for an observable A, if it introduces an interactions with the environment in such a way that the coherent superposition of states is transformed into a mixture of states of the subsystem to be measured.

Despite the fact that we arrive at a statistical mixture of states for our subsystem, there is still a single quantum state for the universe, that is for the measured system, the detector and its environment taken together. However, if we consider this state only in a subsystem, it has the form of a statistical mixture of states from the

Chapter 8

Dynamics

8.1 Dynamics of an expectation value

Here we investigate the time evolution of an expectation value

$$\frac{d}{dt}\langle\psi|\hat{A}|\psi\rangle = \langle\partial_t\psi|\hat{A}|\psi\rangle + \langle\psi|(\frac{d\hat{A}}{dt})|\psi\rangle + \langle\psi|\hat{A}|\partial\psi\rangle$$

We use the time-dependent Schrödinger equation[1] $i\hbar\partial_t|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$ to obtain the derivatives of bras and kets

$$i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle$$
$$-i\hbar\partial_t\langle\psi| = \langle\psi|\hat{H}$$

Thus we obtain

$$\begin{aligned} i\hbar\frac{d}{dt}\langle\psi|\hat{A}|\psi\rangle &= \langle\psi|(-\hat{H}\hat{A} + \hat{A}\hat{H})|\psi\rangle + \langle\psi|i\hbar(\frac{d\hat{A}}{dt})|\psi\rangle \\ &= \langle\psi|\left[\hat{A},\hat{H}\right] + i\hbar\left(\frac{d\hat{A}}{dt}\right)|\psi\rangle \end{aligned} \tag{8.1}$$

Fig. 8.1: Paul Ehrenfest, 1880-1933.

This provides us with an important rule:

COMMUTATOR AND CONSERVATION LAWS

If an observable is not explicitly time dependent and if it commutates with the Hamilton operator, the expectation value of that observable does not change with time.

As a special case we find the energy conservation: the expectation value of the Hamilton operator, the energy, is conserved if the Hamilton operator is not explicitly time dependent.

8.2 Quantum numbers

From the above relation Eq. 8.1 follows the following:

The expectation value of a observable, that commutates with the Hamilton operator is invariant. Such observables are called **constants of motion**.

[1]The rule for the bra is obtained from the rule for the ket by forming a matrix element with an arbitrary bra $\langle\phi|$, and using the rules for interchanging bras and kets in a matrix element, i.e. $\langle\psi|\hat{A}|\phi\rangle = \langle\phi|\hat{A}^\dagger|\psi\rangle$

Definition 8.1 *COMPATIBLE OPERATORS*
Two operators are called compatible if their commutator vanishes.

We see immediately, that for a time-independent Hamilton operator, the energy is a constant of motion, because the Hamilton operator commutates with itself.

If an operator $\hat{A}$ commutates with the Hamilton operator, we can find a basis of states which are eigenstates both of $\hat{A}$ and of $\hat{H}$. The eigenvalues of such an operator $\hat{A}$ are called **good quantum numbers**.

A basis of common eigenstates of $\hat{H}$ and $\hat{A}$, which have a vanishing commutator, is constructed as follows: We start with the eigenstates of the Hamilton operator.

- If an eigenvalue of the Hamiltonian is non-degenerate, the corresponding eigenstate $|\Psi\rangle$ is also an eigenstate of the operator $\hat{A}$. This is shown as follows:

$$\begin{aligned} \hat{H}|\psi\rangle &= |\psi\rangle E \\ \Rightarrow \hat{A}\hat{H}|\psi\rangle &= \hat{A}|\psi\rangle E \\ \stackrel{[H,A]=0}{\Rightarrow} \hat{H}\left(\hat{A}|\psi\rangle\right) &= \left(\hat{A}|\psi\rangle\right) E \end{aligned} \tag{8.2}$$

  Thus, the state $\hat{A}|\psi\rangle$ is also an eigenstate of the Hamilton operator with the same eigenvalue. Because the eigenvalue is non-degenerate, the state $\hat{A}|\psi\rangle$ and the state $|\psi\rangle$ must be linear dependent. Hence

$$\hat{A}|\psi\rangle = |\psi\rangle a$$

  with some scale factor a. Because the above equation, is nothing but an eigenvalue equation, the scale factor a is the corresponding eigenvalue of $\hat{A}$. Thus, the state $|\psi\rangle$ is an eigenstate to the operator $\hat{A}$.

- If the eigenvalue E of the Hamiltonian is m-fold degenerate, we can use one of the eigenstates from the corresponding set of degenerate states to form a sequence of states

$$|\psi_j\rangle = \hat{A}^j|\psi\rangle$$

  which are all eigenstates of the Hamiltonian with the same eigenvalue E. The proof is analogous to the non-degenerate case as shown in Eq. 8.2.

  We determine a linear independent sub-set of the states $|\bar{\psi}_j\rangle$ for example by **Gram-Schmidt orthogonalization**. While there seems to be an infinite number of states $|\psi_j\rangle$ in the sequence, there are only m linear independent eigenstates with this eigenvalue of the Hamilton operator. Thus the linear independent set of states $|\bar{\psi}_j\rangle$ resulting from the Gram-Schmidt orthogonalization has at most m members.

  A state $\hat{A}|\bar{\psi}_j\rangle$ can itself be expressed as a superposition of the states in the set $\left\{|\bar{\psi}_k\rangle; k = 1, \ldots, m\right\}$ of degenerate states.

$$\begin{aligned} \hat{A}|\bar{\psi}_j\rangle &= \sum_k |\bar{\psi}_k\rangle c_{k,j} \\ \Rightarrow \quad \underbrace{\langle\bar{\psi}_i|\hat{A}|\bar{\psi}_j\rangle}_{=A_{i,j}} &= \sum_k \underbrace{\langle\bar{\psi}_i|\bar{\psi}_k\rangle}_{\delta_{i,k}} c_{k,j} = c_{i,j} \end{aligned}$$

  The matrix elements $c_{i,j}$ are equal to the matrix elements $A_{i,j}$ of the operator $\hat{A}$ in the basis $\{|\bar{\psi}_j\rangle\}$. Therefore, we denote it in the following as $A_{i,j} = c_{i,j}$.

  Since the matrix $\boldsymbol{A}$ is hermitian, it can be diagonalized, which yields the eigenvectors $\vec{c}_n$ and eigenvalues a_n.

$$\boldsymbol{A}\vec{c}_n = \vec{c}_n a_n$$

The eigenvectors can then be used to create the eigenstates $|\phi_n\rangle$ of $\hat{A}$

$$|\phi_n\rangle \stackrel{\text{def}}{=} \sum_j |\bar{\psi}_j\rangle c_{j,n}$$

which fulfill the eigenvalue equation

$$\hat{A}|\phi_n\rangle \stackrel{\text{def}}{=} |\phi_n\rangle a_n$$

and which are simultaneously eigenstates of the Hamiltonian.

Thus it has been shown that it is possible to construct a complete orthonormal basisset of states, which are simultaneously eigenstates of $\hat{H}$ and $\hat{A}$, if the two operators commutate.

A complete set of compatible operators is a set of compatible operators, with the property that the set of their eigenvalues completely characterizes each state of a complete basis in Hilbert space.

An example are the states of the Hydrogen atom. They are characterized completely by the three quantum numbers n, ℓ, m through the eigenvalue equations

$$\begin{aligned}
\hat{H}|\psi_{\ell,m,n}\rangle &= |\psi_{\ell,m,n}\rangle \frac{Z}{2n^2}H \\
\hat{L}_z|\psi_{\ell,m,n}\rangle &= |\psi_{\ell,m,n}\rangle \hbar m \\
\hat{\vec{L}}^2|\psi_{\ell,m,n}\rangle &= |\psi_{\ell,m,n}\rangle \hbar^2 \ell(\ell+1)
\end{aligned}$$

where Z is the atomic number, $\hat{H}$ is the Hamilton operator of the hydrogen atom, H is the Hartree energy unit, Z is the atomic number and $\hat{\vec{L}}$ is the angular-momentum operator. n is the principal quantum number, ℓ is the main angular-momentum quantum number and m is the magnetic quantum number.

Another example is the wave vector of a plane wave. Plane waves form a complete set of states. They are eigenstates of the Hamiltonian for a free particle and they are characterized by the eigenvalue equation.

$$\hat{\vec{p}}|\psi_{\vec{k}}\rangle = |\psi_{\vec{k}}\rangle \hbar\vec{k}$$

8.3 Ehrenfest Theorem

If we apply the above rule to the position and momentum operators, we find that the expectation values for these quantities evolve with the classical equations of motion. The Ehrenfest theorem plays an important role to justify the argument that we used in the second chapter to derive the Schrödinger equation. It says, that if a wave packet remains localized, it behaves like a classical particle. We need to read this sentence carefully, though. The wave packet can still split or spread, and then the correspondence between classical mechanics and quantum theory is not as clear cut.

EHRENFEST THEOREM

The **Ehrenfest Theorem**[31] says that the time evolution of a quantum mechanical expectation values is identical to that of a classical variable under the expectation value of the forces.

$$\begin{aligned}
\frac{d}{dt}\langle\psi|\hat{x}|\psi\rangle &= \langle\psi|\partial_p H(\hat{p},\hat{x})|\psi\rangle \\
\frac{d}{dt}\langle\psi|\hat{p}|\psi\rangle &= \langle\psi| -\partial_x H(\hat{p},\hat{x})|\psi\rangle
\end{aligned} \tag{8.3}$$

These equations compare to Hamilton's equations Eq. 3.6 on p. 31

Let us first consider a special Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$$

The general proof will be shown later in this section.

We obtain for the position operator

$$\begin{aligned}
\frac{d}{dt}\langle\psi|\hat{x}|\psi\rangle &\stackrel{\text{Eq. 8.1}}{=} \langle\psi|\frac{i}{\hbar}[\frac{\hat{p}^2}{2m} + V(\hat{x}), \hat{x}]|\psi\rangle = \langle\psi|\frac{i}{2m\hbar}[\hat{p}^2, \hat{x}]|\psi\rangle \\
&= \langle\psi|\frac{i}{2m\hbar}(\hat{p}\hat{p}\hat{x} \underbrace{-\hat{p}\hat{x}\hat{p} + \hat{p}\hat{x}\hat{p}}_{=0} - \hat{x}\hat{p}\hat{p})|\psi\rangle \\
&= \langle\psi|\frac{i}{2m\hbar}(\hat{p}[\hat{p}, \hat{x}] + [\hat{p}, \hat{x}]\hat{p})|\psi\rangle \\
&= \langle\psi|\frac{\hat{p}}{m}|\psi\rangle
\end{aligned}$$

For the momentum operator we find similarly the equation corresponding to the Hamilton equation for the momentum[2].

$$\begin{aligned}
\frac{d}{dt}\langle\psi|\hat{p}|\psi\rangle &\stackrel{\text{Eq. 8.1}}{=} \langle\psi|\frac{i}{\hbar}\left[\frac{\hat{p}^2}{2m} + V(\hat{x}), \hat{p}\right]|\psi\rangle \\
&= \langle\psi|\frac{i}{\hbar}[V(\hat{x}), \hat{p}]|\psi\rangle \\
&\stackrel{\text{Eq. 6.11}}{=} \langle\psi|\left[V(\hat{x}), \underbrace{\int dx\, |x\rangle\partial_x\langle x|}_{\frac{i}{\hbar}\hat{p}}\right]|\psi\rangle \\
&\stackrel{V(\hat{x})|x\rangle=|x\rangle V(x)}{=} \int dx\, \langle\psi|x\rangle\,[V(x), \partial_x]\,\langle x|\psi\rangle \\
&= -\int dx\, \langle\psi|x\rangle\,\underbrace{[\partial_x, V(x)]}_{(\partial_x V)}\langle x|\psi\rangle \\
&= -\langle\psi|(\partial_x V(\hat{x}))|\psi\rangle
\end{aligned}$$

The meaning of $\partial_x V(\hat{x})$ is the following: First form the derivative $(\partial_x V(x))$ of the function $V(x)$. The parenthesis shall indicate that the derivative does not act on any following functions. Then replace the argument x by the operator $\hat{x}$.

We have shown the Ehrenfest theorem for a special form of the Hamilton operator. In order to derive the Ehrenfest theorem in its most general form, we need the equation of motion for operators derived earlier in Eq. 8.1

$$\begin{aligned}
\frac{d}{dt}\langle\psi|\hat{x}|\psi\rangle &\stackrel{\text{Eq. 8.1}}{=} \langle\psi|\frac{i}{\hbar}[\hat{H}, \hat{x}]|\psi\rangle \\
\frac{d}{dt}\langle\psi|\hat{p}|\psi\rangle &\stackrel{\text{Eq. 8.1}}{=} \langle\psi|\frac{i}{\hbar}[\hat{H}, \hat{p}]|\psi\rangle
\end{aligned}$$

and the commutator expressions

$$\begin{aligned}
[H(\hat{p}, \hat{x}), \hat{x}] &= \frac{\hbar}{i}\partial_p H(\hat{p}, \hat{x}) \\
[H(\hat{p}, \hat{x}), \hat{p}] &= -\frac{\hbar}{i}\partial_x H(\hat{p}, \hat{x})
\end{aligned}$$

[2]In the derivation, note that $\partial_x V(x) = (\partial_x V(x)) + V(x)\partial_x$. This is an operator expression that must be read as if it were applied to a state following the expression on the right side.

which are shown as follows:

Proof: Here we show that the operator equation

$$[f(\hat{p},\hat{x}),\hat{x}] = \frac{\hbar}{i}\partial_p f(\hat{p},\hat{x}) \tag{8.4}$$

$$[f(\hat{p},\hat{x}),\hat{p}] = -\frac{\hbar}{i}\partial_x f(\hat{p},\hat{x}) \tag{8.5}$$

is valid for any function of the canonical conjugate operators $\hat{x}$ and $\hat{p}$, that is those having the commutator relation $[\hat{p},\hat{x}] = \frac{\hbar}{i}$

A function of an operator is defined by its power-series expansion. Therefore, any operator function $f(\hat{p},\hat{x}) = \sum_\alpha c_\alpha g_\alpha(\hat{p},\hat{x})$ can be expressed by a sum of weighted powers $g_\alpha(\hat{p},\hat{x}) = \hat{x}^i\hat{p}^j\hat{x}^k\hat{p}^l\ldots$, where $i,j,k,l\ldots$ are non-negative integers. The order of the power is $n_\alpha = i+j+k+l+\ldots$. It is important to remember, that the order of the operators p and x in the function f matters, and this order must be observed when performing the derivatives. In the following, we show first that the above relation holds for arbitrary powers $g_\alpha(\hat{p},\hat{x})$ and than derive from that the equations above for general operator functions $f(\hat{x},\hat{p})$.

The proof for powers $g_\alpha(\hat{p},\hat{x})$ proceeds by induction.

1) We start with $g_\alpha(\hat{p},\hat{x}) = \hat{p}^1$ and $g_{\alpha'}(\hat{p},\hat{x}) = \hat{x}^1$, which are the only powers of order one.

$$[\hat{p},\hat{x}] = \frac{\hbar}{i} = \frac{\hbar}{i}[\partial_p\hat{p}] \tag{8.6}$$

$$[\hat{x},\hat{p}] = -\frac{\hbar}{i} = -\frac{\hbar}{i}[\partial_x\hat{x}] \tag{8.7}$$

Thus, the result is obviously true for powers of order one.

2) Now we assume that the relations Eq. 8.4 and Eq. 8.5 are true for powers of order $n-1$ in $\hat{x}$ and $\hat{p}$ and derive from that the relation for powers of order n. All powers of order n can be written as one of the four possibilities $\hat{x}g(\hat{p},\hat{x})$, $g(\hat{p},\hat{x})\hat{x}$, $\hat{p}g(\hat{p},\hat{x})$, $g(\hat{p},\hat{x})\hat{p}$, where g is a power of $\hat{x}$ and $\hat{p}$ of order $n-1$.

$$\begin{aligned}
[\hat{x}g(\hat{p},\hat{x}),\hat{x}] &= \hat{x}g(\hat{p},\hat{x})\hat{x} - \hat{x}^2 g(\hat{p},\hat{x}) = \hat{x}[g(\hat{p},\hat{x}),\hat{x}] \overset{\text{Eq. 8.4}}{=} \hat{x}\left(\frac{\hbar}{i}\partial_p g(\hat{x},\hat{p})\right)\\
&= \frac{\hbar}{i}\partial_p\Big(\hat{x}g(\hat{x},\hat{p})\Big)\\
[g(\hat{p},\hat{x})\hat{x},\hat{x}] &= g(\hat{p},\hat{x})\hat{x}^2 - \hat{x}g(\hat{p},\hat{x})\hat{x} = [g(\hat{p},\hat{x}),\hat{x}]\hat{x} \overset{\text{Eq. 8.4}}{=} \left(\frac{\hbar}{i}\partial_p g(\hat{x},\hat{p})\right)\hat{x}\\
&= \frac{\hbar}{i}\partial_p\Big(g(\hat{x},\hat{p})\hat{x}\Big)\\
[g(\hat{p},\hat{x})\hat{p},\hat{x}] &= g(\hat{p},\hat{x})\hat{p}\hat{x}\overbrace{-g(\hat{p},\hat{x})\hat{x}\hat{p}+g(\hat{p},\hat{x})\hat{x}\hat{p}}^{=0}-\hat{x}g(\hat{p},\hat{x})\hat{p}\\
&= g(\hat{p},\hat{x})[\hat{p},\hat{x}] + [g(\hat{p},\hat{x}),\hat{x}]\hat{p} \overset{\text{Eq. 8.4}}{=} \frac{\hbar}{i}g(\hat{p},\hat{x}) + \left(\frac{\hbar}{i}\partial_p g(\hat{p},\hat{x})\right)\hat{p}\\
&= \frac{\hbar}{i}\partial_p\Big(g(\hat{p},\hat{x})\hat{p}\Big)\\
[\hat{p}g(\hat{p},\hat{x}),\hat{x}] &= \hat{p}g(\hat{p},\hat{x})\hat{x}\overbrace{-\hat{p}\hat{x}g(\hat{p},\hat{x})+\hat{p}\hat{x}g(\hat{p},\hat{x})}^{=0}-\hat{x}\hat{p}g(\hat{p},\hat{x})\\
&= \hat{p}[g(\hat{p},\hat{x}),\hat{x}] + [\hat{p},\hat{x}]g(\hat{p},\hat{x}) \overset{\text{Eq. 8.4}}{=} \hat{p}\left(\frac{\hbar}{i}\partial_p g(\hat{p},\hat{x})\right) + \frac{\hbar}{i}g(\hat{p},\hat{x})\\
&= \frac{\hbar}{i}\partial_p\Big(\hat{p}g(\hat{p},\hat{x})\Big)
\end{aligned}$$

Thus we have shown that the relation Eq. 8.4 carries over from a power of order n to the next higher order $n+1$.

The proof for Eq. 8.5, the commutator with $\hat{p}$, proceeds analogously. We may even exploit Eq. 8.4 is converted into Eq. 8.5 by an exchange of $\hat{p}$ and $\hat{x}$ and sign-change of $\hbar$.

From the commutator relations[a] follows that

$$[\sum_\alpha c_\alpha g_\alpha(\hat{x},\hat{p}),\hat{x}] = \sum_\alpha c_\alpha[g_\alpha(\hat{x},\hat{p}),\hat{x}]$$

which completes the proof for a general function of the operators $\hat{x}$ and $\hat{p}$. *q.e.d.*

[a] If $\hat{A},\hat{B},\hat{C}$ are operators and c is a number the commutator relations $[c\hat{A},\hat{B}] = c[\hat{A},\hat{B}]$ and $[\hat{A}+\hat{B},\hat{C}] = [\hat{A},\hat{C}] + [\hat{B},\hat{C}]$ hold.

8.4 Particle conservation and probability current

As the system evolves, the probability to find the system anywhere in space does not change. This fundamental conservation law, the particle conservation law, led to the requirement that the Hamiltonian must be Hermitian.[3] Here we investigate the probability for the particle to be at a given point or in a given finite volume respectively. The probability to find the system in a given volume can change only if the particle appears somewhere else. This can be described by the probability that particle flow through the boundary of the volume. In this section we will develop an expression for the probability flux.

The condition that the number of particles in a given volume Ω changes only when particles flow through the surface $\partial\Omega$ of the volume, can be written as follows:

$$\partial_t \int_\Omega d^3r\; P(\vec{r},t) = -\oint_{\partial\Omega} d\vec{A}\; \vec{J}(\vec{r},t)$$

Here $P(\vec{r},t)$ is the probability density and $\vec{J}(\vec{r},t)$ is the probability current. Note that per convention the surface element vector $d\vec{A}$ always points outward. We use Gauss Theorem to formulate this condition in differential form of the conservation law.

$$\begin{aligned}
\int_\Omega d^3r\; \partial_t P(\vec{r},t) &= -\int d^3r\; \vec{\nabla}\vec{J}(\vec{r},t) \\
\int_\Omega d^3r\; \Big(\partial_t P(\vec{r},t) + \vec{\nabla}\vec{J}(\vec{r},t)\Big) &= 0 \\
\partial_t P(\vec{r},t) + \vec{\nabla}\vec{J}(\vec{r},t) &= 0
\end{aligned} \tag{8.8}$$

We have used here that if the integral vanishes for any arbitrary volume Ω, the integrand must vanish anywhere in space.

Let us evaluate the probability of finding the particle in the volume Ω.

$$\begin{aligned}
&\partial_t \int_\Omega d^3r\; \Psi^*(\vec{r},t)\Psi(\vec{r},t) \\
&= \int_\Omega d^3r\; \Big[\Big(\partial_t\Psi^*(\vec{r},t)\Big)\Psi(\vec{r},t) + \Psi^*(\vec{r},t)\Big(\partial_t\Psi(\vec{r},t)\Big)\Big] \\
&= \frac{i}{\hbar}\int_\Omega d^3r\; \Big[\Psi(\vec{r},t)\Big(-\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r})\Big)\Psi^*(\vec{r},t)) - \Psi^*(\vec{r},t)\Big(-\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r})\Big)\Psi(\vec{r},t)\Big] \\
&= \frac{-i\hbar}{2m}\int_\Omega d^3r\; \Big[\Psi(\vec{r},t)\vec{\nabla}^2\Psi^*(\vec{r},t) - \Psi^*(\vec{r},t)\vec{\nabla}^2\Psi(\vec{r},t)\Big] \\
&= \frac{-i\hbar}{2m}\int_\Omega d^3r\; \Big[\vec{\nabla}\Big(\Psi(\vec{r},t)\vec{\nabla}\Psi^*(\vec{r},t)\Big) - \Big(\vec{\nabla}\Psi(\vec{r},t)\Big)\vec{\nabla}\Psi^*(\vec{r},t) \\
&\quad -\vec{\nabla}\Big(\Psi^*(\vec{r},t)\vec{\nabla}\Psi(\vec{r},t)\Big) + \Big(\vec{\nabla}\Psi^*(\vec{r},t)\Big)\vec{\nabla}\Psi(\vec{r},t)\Big] \\
&= \frac{-i\hbar}{2m}\int_\Omega d^3r\; \vec{\nabla}\Big[\Psi(\vec{r},t)\vec{\nabla}\Psi^*(\vec{r},t) - \Psi^*(\vec{r},t)\vec{\nabla}\Psi(\vec{r},t)\Big]
\end{aligned}$$

[3] Particle conservation requires that $\partial_t\langle\psi|\psi\rangle = 0$.

$$0 = \partial_t\langle\psi|\psi\rangle = \langle\partial_t\psi|\psi\rangle + \langle\psi|\partial_t\psi\rangle$$

Now we use the Schrödinger equation

$$i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle \Rightarrow |\partial_t\psi\rangle = -\frac{i}{\hbar}\hat{H}|\psi\rangle \qquad \Rightarrow \qquad \langle\partial_t\psi| = \frac{i}{\hbar}\langle\partial_t\psi|\hat{H}^\dagger$$

and insert into the above equation

$$0 = \langle\partial_t\psi|\psi\rangle + \langle\psi|\partial_t\psi\rangle = \frac{i}{\hbar}\Big(\langle\psi|\hat{H}^\dagger|\psi\rangle - \langle\psi|\hat{H}|\psi\rangle\Big) = \frac{i}{\hbar}\langle\psi|\hat{H}^\dagger - \hat{H}|\psi\rangle$$

Since this equation holds for any wave function $|\psi\rangle$ it follows that $\hat{H}^\dagger - \hat{H}$ has only zero eigenvalues and therefore is the zero operator. hence $\hat{H}^\dagger = \hat{H}$, which is the condition that the Hamiltonian is hermitian.

Let us define[4] now the **current density**

PROBABILITY-CURRENT DENSITY

$$\vec{J}(\vec{r},t)=\frac{i\hbar}{2m}\Big[\Psi(\vec{r},t)\vec{\nabla}\Psi^*(\vec{r},t)-\Psi^*(\vec{r},t)\vec{\nabla}\Psi(\vec{r},t)\Big]=\mathrm{Re}\Big[\Psi^*(\vec{r,t})\frac{\hat{\vec{p}}}{m_0}\psi(\vec{r},t)\Big] \tag{8.9}$$

Note that the current density has a different functional form for other Hamiltonians.[5]

The current density could also have been derived from the Noether theorem as a consequence of **gauge symmetry**. Gauge symmetry implies that a multiplication of the wave functions with a complex phase factor $e^{i\phi}$ with a real angle ϕ does not change the physics. More precisely stated, the action is invariant with respect to a multiplication with a phase factor. The Noether theorem is applied for the Klein-Gordon field in ΦSX:Elektrodynamik[32].

8.5 Schrödinger, Heisenberg and Interaction pictures

8.5.1 Schrödinger picture

We will learn several representations of wave functions and operators. These are the Heisenberg picture and the Interaction picture. In contrast, the conventional picture is the **Schrödinger picture**. In the Schrödinger picture the operators are time independent, except for an explicit time dependence such as a time dependent potential, and the dynamics of the system is described by a differential equation for the wave function, the Schrödinger equation.

$$i\hbar\partial_t|\psi_S(t)\rangle=\hat{H}_S|\psi_S(t)\rangle$$

The subscript S indicates that the wave functions and operators are given in the conventional picture, the schrödinger equation. The expectation value is given by

$$\langle A\rangle=\langle\psi_S|\hat{A}_S|\psi_S\rangle$$

8.5.2 Heisenberg picture

In order to introduce the different pictures, we will require the concept of a so-called **propagator** [6].

In this section we limit ourselves to time-independent Hamiltonians. The more general case is discussed in section J on p. 301.

Formally, we can directly solve the Schrödinger equation for a time-independent Hamiltonian in the following way

$$|\psi_S(t)\rangle=\hat{U}(t)|\psi_S(0)\rangle. \tag{8.10}$$

where the propagator fulfills the equation

$$i\hbar\partial_t\hat{U}(t)=\hat{H}_S\hat{U}(t) \tag{8.11}$$

which is proven by inserting Eq. 8.10 into the time-dependent Schrödinger equation.

[4]This is actually the definition of the particle current.The usual approach is to convert the time derivative of the Probability density into the divergence of a vector field. If this is successful, we can claim that the particle number is conserved and the vector field is the corresponding current density of those particles.

[5]For a particle in an electric and magnetic field, so that $H=(\vec{p}-q\vec{A}(r))^2/2m+e\Phi(r)$, the current density has the form $\vec{J}=\frac{i\hbar}{2m}\Big[\Psi(r,t)\vec{\nabla}\Psi^*(r,t)-\Psi^*(r,t)\vec{\nabla}\Psi(r,t)\Big]-\frac{e}{m}\vec{A}\Psi^*\Psi$. For massless particles with $H=\vec{e}cp$ such as photons the current is $J=\vec{e}cP$, where $\vec{e}$ is the direction of the light beam.

[6]"propagator" translates as "Zeitentwicklungsoperator" into German

We obtain the propagator as

$$\hat{U}(t) = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}_S t}$$

which is proven by insertion into the defining equation Eq. 8.11.

The propagator is unitary[7], that is

$$\hat{U}(-t) = \hat{U}^{-1}(t) = \hat{U}^\dagger(t)$$

The expectation value of an operator at a particular time can be written in two different ways, once using the state at that time, and once the state at a reference time 0.

$$\langle A\rangle = \langle\psi_S(t)|\hat{A}_S|\psi_S(t)\rangle = \langle\psi_S(0)|\hat{U}^\dagger(t)\hat{A}_S\hat{U}(t)|\psi_S(0)\rangle$$

The operator

$$\hat{A}_H(t) \stackrel{\text{def}}{=} \hat{U}^\dagger(t)\hat{A}_S(t)\hat{U}(t)$$

is the corresponding **Heisenberg operator** and the **Heisenberg state** is

$$|\Psi_H\rangle \stackrel{\text{def}}{=} |\Psi_S(0)\rangle = \hat{U}^\dagger(t)|\Psi_S(t)\rangle$$

As the Heisenberg state is time-independent, the dynamics is entirely contained now in the operators. We can write down the equations of motion for this operator as [8]

$$i\hbar\partial_t\hat{A}_H(t) = [\hat{A}_H(t), \hat{H}_S] + i\hbar(\frac{\partial A}{\partial t})_H$$

Note here that the last term means $(\frac{\partial A}{\partial t})_H \stackrel{\text{def}}{=} \hat{U}^\dagger\frac{d\hat{A}_S}{dt}\hat{U}$.

If $\hat{H}$ is time independent, the transformation can be written explicitly in terms of eigenvalues ϵ_n and eigenstates $|\phi_n\rangle$of the reference Hamiltonian.

$$\hat{U}(t) = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}_S t} = \sum_n |\phi_n\rangle\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n t}\langle\phi_n|$$

so that

$$\hat{A}_H = \sum_{n,m}|\phi_n\rangle\langle\phi_n|\hat{A}_S|\phi_m\rangle\mathrm{e}^{\frac{i}{\hbar}(\epsilon_n-\epsilon_m)t}\langle\phi_m|$$

[7] without proof

[8] In the derivation we use that $\hat{H}$ commutates with the propagator $\hat{U}$, and that the time derivative of the propagator is $\frac{i}{\hbar}\hat{H}\hat{U}$.

$$\begin{aligned}
\frac{d}{dt}\hat{A}_H &= \frac{d}{dt}\left[\hat{U}^\dagger(t)\hat{A}_S\hat{U}(t)\right]\\
&= \left(\frac{d}{dt}\hat{U}^\dagger(t)\right)\hat{A}_S\hat{U}(t) + \hat{U}^\dagger(t)\left(\frac{d}{dt}\hat{A}_S\right)\hat{U}(t) + \hat{U}^\dagger(t)\hat{A}_S\left(\frac{d}{dt}\hat{U}\right)\\
&= \left(\frac{i}{\hbar}\hat{H}_S\hat{U}^\dagger(t)\right)\hat{A}_S\hat{U}(t) + \hat{U}^\dagger(t)\left(\frac{d}{dt}\hat{A}_S\right)\hat{U}(t) + \hat{U}^\dagger(t)\hat{A}_S\left(-\frac{i}{\hbar}\hat{U}\hat{H}\right)\\
&= \frac{i}{\hbar}\left[\hat{H}(t)\hat{U}^\dagger(t)\hat{A}_S\hat{U}(t) - \hat{U}^\dagger(t)\hat{A}_S\hat{U}(t)\hat{H}\right] + \hat{U}^\dagger(t)\left(\frac{d}{dt}\hat{A}_S\right)\hat{U}(t)\\
&= \frac{i}{\hbar}\left[\hat{H}(t), \hat{A}_H\right] + (\frac{\partial\hat{A}}{\partial t})_H
\end{aligned}$$

8.5.3 Interaction picture

Often we can divide the Hamiltonian into a Hamiltonian H_0 of a reference system, for which the Schrödinger equation can be solved exactly, and a remainder W. W is often called the interaction.

$$\begin{aligned}\hat{H}_S &= \hat{H}_{0,S} + \hat{W}_S \\ i\hbar\partial_t|\psi_S(t)\rangle &= \Big(\hat{H}_{0,S} + \hat{W}_S\Big)|\psi_S(t)\rangle\end{aligned}$$

Let us now introduce the propagator $\hat{U}_0(t)$ of the unperturbed system

$$i\hbar\partial_t\hat{U}_0(t) = \hat{H}_S\hat{U}_0(t) \tag{8.12}$$

For a time-independent Hamiltonian we can integrate this equation directly and obtain

$$\hat{U}_0(t) = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}_{0,S}t}$$

Now we introduce a state in the interaction picture

$$|\Psi_W\rangle \stackrel{\text{def}}{=} \hat{U}_0^\dagger(t)|\Psi_S(t)\rangle \tag{8.13}$$

and an operator in the interaction picture

$$\hat{A}_W(t) \stackrel{\text{def}}{=} \hat{U}_0^\dagger(t)\hat{A}_S\hat{U}_0(t) \tag{8.14}$$

We observe that the state in the interaction picture is just the initial state in the absence of a perturbation $\hat{W}$. If the perturbation is small also the time dependence is slow. This is the main reason to introduce the interaction picture, namely to work with states for which the dynamics is slow and governed only by the "difficult" part of the Hamiltonian.

We can show that expectation values have the usual form also in the interaction picture, namely

$$\langle A\rangle = \langle\psi_S|\hat{A}_S|\psi_S\rangle \stackrel{\hat{U}_0^\dagger\hat{U}_0=\hat{1}}{=} \langle\psi_S|\underbrace{\hat{U}_0\hat{U}_0^\dagger}_{\hat{1}}\hat{A}_S\underbrace{\hat{U}_0\hat{U}_0^\dagger}_{\hat{1}}|\psi_S\rangle \stackrel{\text{Eqs. 8.13,8.14}}{=} \langle\psi_W|\hat{A}_W|\psi_W\rangle$$

Let us now investigate the equation of motion for a state in the Interaction picture

$$\begin{aligned} & & i\hbar\partial_t|\psi_S(t)\rangle &= \Big(\hat{H}_{0,S} + \hat{W}_S\Big)|\psi_S(t)\rangle \\ &\stackrel{\text{Eq. 8.13}}{\Rightarrow} & i\hbar\partial_t\hat{U}_0(t)|\psi_W(t)\rangle &= \Big(\hat{H}_{0,S} + \hat{W}\Big)\hat{U}_0(t)|\psi_W(t)\rangle \\ \stackrel{\text{Eq. 8.12}}{\Rightarrow} & & \hat{U}_0(t)\Big(H_0 + i\hbar\partial_t\Big)|\psi_W(t)\rangle &= \Big(\hat{H}_{0,S} + \hat{W}_S\Big)\hat{U}_0(t)|\psi_W(t)\rangle \\ &\stackrel{[\hat{H}_{0,S}\hat{U}_0]=0}{\Rightarrow} & i\hbar\partial_t|\psi_W(t)\rangle &= \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0t}\hat{W}_S\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0t}|\psi_W(t)\rangle = \hat{W}_W|\psi_W(t)\rangle\end{aligned}$$

Hence

$$i\hbar\partial_t|\psi_W(t)\rangle = \hat{W}_W|\psi_W(t)\rangle \tag{8.15}$$

The dynamics for the operator is obtained as follows

$$
\begin{aligned}
i\hbar\partial_t\hat{A}_W &= i\hbar\partial_t\left(\hat{U}_0^\dagger\hat{A}_S\hat{U}_0\right)\\
&= \underbrace{\left(i\hbar\partial_t\hat{U}_0^\dagger\right)}_{-\left(i\hbar\partial_t\hat{U}_0\right)^\dagger}\hat{A}_S\hat{U}_0+\hat{U}_0^\dagger\left(i\hbar\partial_t\hat{A}_S\right)\hat{U}_0+\hat{U}_0^\dagger\hat{A}_S\left(i\hbar\partial_t\hat{U}_0\right)\\
&= -\underbrace{\left(\hat{H}_{0,S}\hat{U}_0\right)^\dagger}_{\hat{H}_{0,S}\hat{U}_0^\dagger}\underbrace{\hat{U}_0\hat{U}_0^\dagger}_{=\hat{1}}\hat{A}_S\hat{U}_0+i\hbar\hat{U}_0^\dagger\left(\partial_t\hat{A}_S\right)\hat{U}_0+\hat{U}_0^\dagger\hat{A}_S\underbrace{\hat{U}_0\hat{U}_0^\dagger}_{=\hat{1}}\left(\hat{H}_{0,S}\hat{U}_0\right)\\
&= [\hat{A}_W,\hat{H}_{0,W}]+i\hbar\left(\partial_t\hat{A}\right)_W\\
&\overset{[\hat{H}_0,\hat{U}_0]=0}{=} [\hat{A}_W,\hat{H}_{0,S}]+i\hbar\left(\partial_t\hat{A}\right)_W
\end{aligned}
\tag{8.16}
$$

Thus we see the the operators are propagated via the simple dynamics that is the dynamics is governed by the unperturbed system.

We will come back to the interaction picture later, in section 13.4.1 on p. 205.

Chapter 9

Symmetry

9.1 Introduction

Symmetry is one of the most fundamental and useful concepts in physics. This is so, because an underlying symmetry can often be recognized by inspection. We are biologically prepared by our mental and sensory system to recognize symmetries and their violation.

Our sensitivity towards symmetries has probably developed evolutionary from the necessity to detect, for example, a predator in the high-growing grass of the Savanna.

Human art is largely about symmetries, i.e. order, and the violation of expected symmetries. Symmetry or the absence of symmetry are concepts that have a special access to our mind and emotions. Music, for example, is nearly unconceivable without some tonal and rhythmic systems. Free jazz is an extreme example of destroying such order to a degree that it is nearly unrecognizable. The inexperienced listener is appalled, but the trained ear is fascinated and stimulated by the mental act of recognizing the barely noticeable ordering principles, that are more implicit than explicit.

Fig. 9.1: Wolfgang Pauli, 1900-1958. Austrian Physicist. Nobel price in physics 1945 for the Pauli-exclusion principle.

We say that a system is symmetric under an operation $\hat{S}$, if it behaves in exactly the same way, after all relevant parts are transformed by the same transformation. When we speak about the behavior, we mean the equations of motion. In quantum mechanics the dynamics is governed by the Hamilton operator.

Thus we say that the system is symmetric under an operation $\hat{S}$, if every solution of Hamilton operator is transformed by the operation $\hat{S}$ into a solution of the Hamilton operator.
To be more concise, the system described by the Hamilton operator $\hat{H}$ is **symmetric** under the operation $\hat{S}$, if for any solution of

$$i\hbar\partial_t|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$

the transformed solution

$$|\psi'(t)\rangle \stackrel{\text{def}}{=} \hat{S}|\psi(t)\rangle$$

fulfills the equation

$$i\hbar\partial_t|\psi'(t)\rangle = \hat{H}|\psi'(t)\rangle$$

If the Hamilton operator $\hat{H}$ is symmetric under $\hat{S}$, we call $\hat{S}$ a **symmetry operation** of $\hat{H}$.

Symmetry plays a very important role in quantum theory, because it allows us to break up the problem of solving the Schrödinger equation into smaller problems of reduced dimensions. Since the difficulty to solve a quantum mechanical system is substantial, the use of symmetry is very important. One may object that symmetry is often broken and that the number of systems obeying a symmetry is very limited. This may be so, but symmetry is useful even if the symmetry is not exactly obeyed. We will learn in the section about approximations, that we can start solving a system accurately, even though approximately, if we can find only a similar system that can be solved exactly. Here we can use a symmetric model, even though the symmetry only holds approximately in the real system.

The most simple symmetry operation is the inversion in one dimension $\hat{S} = \int dx\, |-x\rangle\langle x|$. If we apply the inversion to a state $|\psi\rangle$ we obtain the new state $|\psi'\rangle \stackrel{\text{def}}{=} \hat{S}|\psi\rangle$. The wave function is given by

$$\begin{aligned}\psi'(x) &= \langle x|\psi'\rangle = \langle x|\hat{S}|\psi\rangle = \langle x|\int dx'\, |-x'\rangle\langle x'|\psi\rangle \\ &= \int dx'\, \underbrace{\langle x|-x'\rangle}_{\delta(x+x')}\underbrace{\langle x'|\psi\rangle}_{\psi(x')} = \langle -x|\psi\rangle = \psi(-x)\end{aligned}$$

The most important symmetry operations are translations and rotations. In order to grasp the general idea of a symmetry operation, let us explore symmetry operations by collecting some general features that all symmetry operations must obey.

- .

  All symmetry operations have in common that they transform a physical state again into a physical state. This implies that the total probability remains unchanged by a symmetry operation $\hat{S}$.

  Thus, if the state $|\phi\rangle$ is obtained from an arbitrary state $|\psi\rangle$ by a symmetry operation $\hat{S}$, i.e.

  $$|\phi\rangle = \hat{S}|\psi\rangle$$

  then the norm of $|\phi\rangle$ must be identical to that of the original state $|\phi\rangle$, that is

  $$\begin{aligned}\langle\phi|\phi\rangle &= \langle\psi|\psi\rangle \\ \Rightarrow \langle\phi|\phi\rangle &= \langle\hat{S}\psi|\hat{S}\psi\rangle = \langle\psi|\hat{S}^\dagger\hat{S}|\psi\rangle = \langle\psi|\hat{1}|\psi\rangle \\ &\stackrel{\text{true for all } |\psi\rangle}{\Rightarrow} \hat{S}^\dagger\hat{S} = \hat{1}\end{aligned}$$

In the last step we exploited that two operators are identical, if the expectation values of both operators are identical for all states.

This implies that a symmetry transformation is represented by a unitary operator S, i.e.

$$\hat{S}^\dagger\hat{S} = \hat{1} \qquad \text{or} \qquad \hat{S}^{-1} = \hat{S}^\dagger. \tag{9.1}$$

- An operator $\hat{A}$ is symmetrical under the transformation $\hat{S}$, if its expectation values are unchanged for each state when applying the transformation[1]

$$|\phi\rangle = \hat{S}|\psi\rangle \qquad \Rightarrow \qquad \langle\phi|\hat{A}|\phi\rangle = \langle\psi|\hat{A}|\psi\rangle$$

From this identity we find that $\hat{S}^\dagger\hat{A}\hat{S} = \hat{A}$. Applying $\hat{S} = (\hat{S}^\dagger)^{-1}$ to the left on both sides of the equation shows the following

An operator $\hat{A}$ is symmetric under a transformation $\hat{S}$, if it commutates with the transformation, i.e.

$$[\hat{A}, \hat{S}] = 0 \tag{9.2}$$

We have already seen that if two operators commutate, we can find a complete system of states, that are eigenstates to both operators.

- .

A system is called symmetric with respect to a transformation, when the Hamilton operator is symmetric.

Now we can come to the main point of symmetry operations: the knowledge of symmetries can tremendously simplify the solution of the Schrödinger equation. We will see in the following, that it is fairly easy to determine the eigenstates of a symmetry operator. We have seen that we can determine the eigenstates of the Hamilton operator from the eigenstates of the symmetry operator.

In a basisset of eigenstates of a symmetry operator $\hat{S}$, the Hamilton matrix is block diagonal. That is, for each pair of eigenstates of $\hat{S}$ with different eigenvalues s_1, s_2, i.e.

$$\hat{S}|\psi_1\rangle = |\psi_1\rangle s_1 \qquad \wedge \qquad \hat{S}|\psi_2\rangle = |\psi_2\rangle s_2$$

the Hamilton matrix element vanishes

$$\langle\psi_1|\hat{H}|\psi_2\rangle = 0 \qquad \text{if} \qquad s_1 \neq s_2 \tag{9.3}$$

The Hamilton matrix is block diagonal in the basis of eigenstates of the symmetry operator.

[1]As an example imagine an operator that describes a spherical potential and the rotation as symmetry operation. Clearly if we rotate a state about the center of the potential, the matrix elements are not affected by the potential.

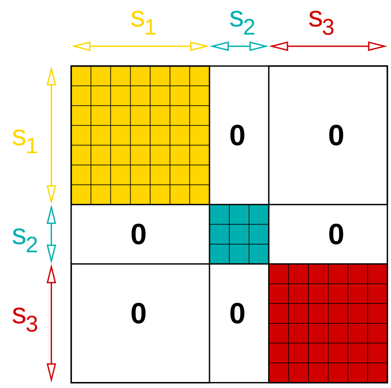


Fig. 9.2: If the Hamilton operator commutates with a symmetry operator $\hat{S}$, the matrix of the Hamilton operator in a basisset of eigenstates of the Symmetry operator $\hat{S}$, is block-diagonal. This means that the matrix elements between eigenstates with different eigenvalues of $\hat{S}$ vanish. As a consequence one only needs to consider, for each eigenvalue of $\hat{S}$ a smaller problem with less basis-functions. Each block typically still contains an infinite number of eigenstates. However, the spectrum of the Hamilton operator for each block is less dense, and this is what matters.

PROOF: If we consider two eigenstates of a symmetry operator with different eigenvalues, say

$$\hat{S}|\psi_1\rangle = |\psi_1\rangle s_1$$
$$\hat{S}|\psi_2\rangle = |\psi_2\rangle s_2$$

we can immediately show that the Hamilton matrix elements between the two states vanishes.
In the derivation, following below, we will use

$$\hat{S}^\dagger|\psi_1\rangle = |\psi_1\rangle s_1^*$$

which is obtained as follows: We exploit Eq. 9.1 saying that the symmetry transformation is unitary, and Eq. 9.4, specified below, saying that the absolute value of all eigenvalues of a unitary operator is equal to one.

$$\hat{S}^\dagger|\psi_1\rangle \stackrel{\hat{S}\hat{S}^\dagger=\hat{1}}{=} \hat{S}^{-1}|\psi_1\rangle \stackrel{\hat{S}|\psi_1\rangle=|\psi_1\rangle s_1}{=} |\psi_1\rangle s_1^{-1} = |\psi_1\rangle \frac{s_1^*}{s_1 s_1^*} \stackrel{\hat{S}\hat{S}^\dagger=\hat{1}}{=} |\psi_1\rangle s_1^*$$

Now we are ready for the actual proof:

$$\begin{aligned}
\langle\psi_1|\hat{H}|\psi_2\rangle s_2 &\stackrel{\hat{S}|\psi_2\rangle=|\psi_2\rangle s_2}{=} \langle\psi_1|\hat{H}\hat{S}|\psi_2\rangle \stackrel{[\hat{H},\hat{S}]=0}{=} \langle\psi_1|\hat{S}\hat{H}|\psi_2\rangle \\
&= \Big(\langle\psi_2|\overbrace{\hat{H}^\dagger\hat{S}^\dagger}^{(\hat{S}\hat{H})^\dagger}|\psi_1\rangle\Big)^* \\
&\stackrel{\hat{S}^\dagger|\psi_1\rangle=|\psi_1\rangle s_1^*}{=} \Big(\langle\psi_2|\hat{H}^\dagger|\psi_1\rangle s_1^*\Big)^* \\
&= \langle\psi_1|\hat{H}|\psi_2\rangle s_1
\end{aligned}$$

Thus

$$\Rightarrow \langle\psi_1|\hat{H}|\psi_2\rangle(s_1 - s_2) = 0$$

Thus if $s_1 \neq s_2$, the matrix element $\langle\psi_1|\hat{H}|\psi_2\rangle$ vanishes. *q.e.d*

9.1.1 Some properties of unitary operators

Unitary operators have a special property that will be useful in the following:

The absolute value of all eigenvalues of a unitary operator is one.

$$\hat{S}\hat{S}^\dagger = \hat{1} \qquad \wedge \qquad \hat{S}|\psi\rangle = |\psi\rangle s \qquad \Rightarrow \qquad |s| = 1 \tag{9.4}$$

Hence we can write its eigenvalues in the form $e^{i\gamma}$ with a real phase-angle γ.

PROOF: We use $\langle\psi|\hat{S}^\dagger\hat{S}|\psi\rangle = \langle\psi|\psi\rangle$ with an eigenstate $|\psi\rangle$ of $\hat{S}$ to an eigenvalue α, i.e. $\hat{S}|\psi\rangle = |\psi\rangle\alpha$. We find $\alpha^*\langle\psi|\psi\rangle\alpha = \langle\psi|\psi\rangle$ and therefore $\alpha^*\alpha = 1$. The most general form for α fulfilling this condition is $\alpha = e^{i\gamma}$ with a real γ. *q.e.d.*

9.1.2 Symmetry groups

A group $\mathcal{G}$ is a set of elements $\hat{S}_n$ with a binary operation (we will refer to it as multiplication, even though it may be a generalization of a multiplication) such that

- the product of two elements of the group is again an element of the group, that is

$$\forall_{\hat{S}_n,\hat{S}_m\in\mathcal{G}} \qquad \hat{S}_n\hat{S}_m \in \mathcal{G}$$

- there is a unity element 1 in the group so that

$$\forall_{\hat{S}_n\in\mathcal{G}} \qquad \hat{S}_n 1 = 1\hat{S}_n = \hat{S}_n$$

- for each element $\hat{S}_n$ in the group there is an inverse $\hat{S}_n^{-1}$ so that

$$\forall_{\hat{S}_n\in\mathcal{G}}\exists_{\hat{S}_n^{-1}\in\mathcal{G}} \qquad \hat{S}_n\hat{S}_n^{-1} = \hat{S}_n^{-1}\hat{S}_n = 1$$

For example, the set of integers with the addition is a group. In this case the addition plays the role of the operation in the group definition. The addition by zero plays the role of the multiplication by the unit element.

It is important to remember that the multiplication does not need to commutate. If the multiplication commutates, $\forall_{\hat{S}_n\hat{S}_m\in\mathcal{G}}\hat{S}_n\hat{S}_m = \hat{S}_m\hat{S}_n$, then the group is called **Abelian**, otherwise it is called **non-Abelian**.

A group is called **finite**, if it has a finite number of elements.

9.2 Finite symmetry groups

First, we consider finite symmetry groups, because they are simple to handle. An example for a finite group is the mirror reflection.

The eigenvalues s_j of a symmetry operator $\hat{S}$ of a finite symmetry group, that is $\hat{S}^n = \hat{1}$ with some positive integer n, are

$$s_j = e^{i2\pi j/n} \qquad \text{with } j \in \{0, 1, \dots, n-1\} \tag{9.5}$$

PROOF: Let us consider a series of operators in the group $(\ldots,\hat{S}^{-1},\hat{1},\hat{S},\hat{S}^2,\ldots)$ If the group is finite, this series must be cyclic, i.e. we can find a number n so that $\hat{S}^{j+n}=\hat{S}^j$ for all j. n must be smaller than or equal to the number of elements in the group. Thus we find, by applying $\hat{S}^{-j}$, that $\hat{S}^n=1$.
Let us consider an eigenstate $|\psi\rangle$ of $\hat{S}$, so that

$$\hat{S}|\psi\rangle=|\psi\rangle e^{i\phi}$$

We have shown that the eigenvalues of a unitary operator can be written as $e^{i\phi}$. Then we can show that $|\psi\rangle$ is an eigenstate of S^j with eigenvalue $e^{i\phi j}$.

$$\hat{S}^j|\psi\rangle=|\psi\rangle e^{i\phi j}$$

Since on the other hand $S^n=1$ we find that $e^{i\phi n}=1$ and therefore $\phi=2\pi/n$. Thus the eigenvalues are $e^{i2\pi j/n}$. *q.e.d.*

One can construct eigenstates of the symmetry operator $\hat{S}$ from any state $|\phi\rangle$ by

$$|\psi_k\rangle=\sum_{j=0}^{n-1}\hat{S}^j|\phi\rangle e^{-i2\pi jk/n} \tag{9.6}$$

Note however, that the result of the sum may vanish.

Here we show[a] that the state defined in Eq. 9.6 is an eigenstate of $\hat{S}$ with eigenvalue $e^{i2\pi k/n}$, where the series $\hat{S}^i$ is cyclic with period n.
Proof: We multiply $|\psi_k\rangle$ from Eq. 9.6 with $\hat{S}$

$$\hat{S}\underbrace{\sum_{j=0}^{n-1}\hat{S}^j|\phi\rangle e^{-i2\pi jk/n}}_{|\psi_k\rangle}=\left(\sum_{j=0}^{n-1}\hat{S}^{j+1}|\phi\rangle e^{-i2\pi(j+1)k/n}\right)e^{i2\pi k/n}=\underbrace{\left(\sum_{j=0}^{n-1}\hat{S}^j|\phi\rangle e^{-i2\pi jk/n}\right)}_{|\psi_k\rangle}e^{i2\pi k/n}$$

In the last step, we shifted the indices $j+1\rightarrow j$ and exploited that the series (S^i) is cyclic with period n. Thus we have shown that $|\psi_k\rangle$ as defined in Eq. 9.6 obeys the eigenvalue equation.

$$\hat{S}|\psi_k\rangle=|\psi_k\rangle e^{i2\pi k/n}$$

q.e.d.

[a] One can *derive* the result by inserting an Ansatz $|\psi_k\rangle=\sum_{j=0}^{n-1}\hat{S}^j|\phi\rangle c_{j,k}$ into the eigenvalue equation $\hat{S}|\psi_k\rangle=|\psi_k\rangle e^{-i2\pi jk/n}$ which results in a recursive relation for the coefficients $c_{j,k}$.

9.3 Continuous symmetries

9.3.1 Shift operator

Let us now consider an operator Q and its canonical momentum P. The canonical momentum has been defined in Eq. 5.26 by

$$[\hat{P},\hat{Q}]\stackrel{\text{Eq. 5.26}}{=}\frac{\hbar}{i}.$$

We define a **shift operator**

SHIFT OPERATOR

$$\hat{S}(\Delta) = \mathrm{e}^{-\frac{i}{\hbar}\hat{P}\Delta}. \tag{9.7}$$

As shown below, the shift operator has the property

$$\hat{S}^\dagger(\Delta)\hat{Q}\hat{S}(\Delta) = \hat{Q} + \Delta\hat{1}. \tag{9.8}$$

This implies that if we transform a state $|\psi\rangle$ to a new state $|\phi\rangle = S(\Delta)|\psi\rangle$, the expectation value of the new state is that of the old state shifted by Δ if the state is normalized, i.e. $\langle\phi|\hat{Q}|\phi\rangle = \langle\psi|\hat{Q}|\psi\rangle + \Delta$.

Here we prove Eq. 9.8 for a shift operator defined by Eq. 9.7.
PROOF: First, we show that the commutator of a function of a momentum with the position can be expressed by the derivative of that function with respect to the momentum, i.e.

$$[f(\hat{P}), \hat{Q}] = \frac{\hbar}{i}\frac{d}{dP}f(\hat{P}) \tag{9.9}$$

Because the function of an operator is defined by its power series expansion, we show first that the relation holds for all powers of the momentum,

$$\begin{aligned}[\hat{P}^n, \hat{Q}] &= \hat{P}^n\hat{Q} - \hat{Q}\hat{P}^n = \hat{P}^{n-1}[\hat{P}, \hat{Q}] + \hat{P}^{n-1}\hat{Q}\hat{P} - \hat{Q}\hat{P}^n \\ &= \frac{\hbar}{i}\hat{P}^{n-1} + [\hat{P}^{n-1}, \hat{Q}]\hat{P}\end{aligned}$$

We apply this equation recursively n times until we obtain on the right side the commutator $[\hat{P}^0, \hat{Q}] = [\hat{1}, \hat{Q}] = \hat{\emptyset}$. Thus we obtain

$$[\hat{P}^n, \hat{Q}] = \frac{\hbar}{i}n\hat{P}^{n-1},$$

which we can rewrite in the form

$$[\hat{P}^n, \hat{Q}] = \frac{\hbar}{i}\frac{d}{dP}\hat{P}^n. \tag{9.10}$$

which is exactly Eq. 9.9 specialized to powers, i.e. $f(x) = x^n$. Addition and multiplication with scalar does not affect the result so that we can generalize Eq. 9.10 to a power-series expansion an such to Eq. 9.9.
This is the first time that we see a derivative with respect to an operator. It is the elementary step *defining* the derivative with respect to an operator. The derivative is formed here just as if $\hat{P}$ would be a simple number. If the function depends on operators that do not commute, the derivative is not as simple and must be traced back to the individual terms of the power series expansion.
Finally, we only need to specialize the result to the shift operator from Eq. 9.7.

$$\begin{aligned}\Rightarrow [\hat{S}(\Delta), \hat{Q}] &\overset{\text{Eq. 9.9}}{=} \frac{\hbar}{i}\frac{d}{dP}\mathrm{e}^{-\frac{i}{\hbar}\hat{P}\Delta} = -\Delta\underbrace{\hat{S}(\Delta)}_{\mathrm{e}^{-\frac{i}{\hbar}\hat{P}\Delta}} \\ \Rightarrow \hat{Q}\hat{S}(\Delta) &= \hat{S}(\Delta)\hat{Q} + \Delta\hat{S}(\Delta) \\ \overset{S^{-1}=S^\dagger}{\Rightarrow} \quad \hat{S}^\dagger(\Delta)\hat{Q}\hat{S}(\Delta) &= \hat{Q} + \Delta\hat{1}\end{aligned}$$

q.e.d. which proves Eq. 9.8.

This shows that if we have an eigenstate of the operator $\hat{Q}$, the shift operator, Eq. 9.7, creates a new eigenstate of $\hat{Q}$ with an eigenvalue increased by Δ.

$$\begin{aligned}&\hat{Q}|\psi\rangle = |\psi\rangle q \\ &\Rightarrow \hat{Q}\Big(\hat{S}(\Delta)|\psi\rangle\Big) = \hat{S}(\Delta)\Big(\hat{Q} + \Delta\Big)|\psi\rangle = \Big(\hat{S}(\Delta)|\psi\rangle\Big)(q + \Delta)\end{aligned}$$

So far we have started from the canonical momentum $\hat{P}$ and constructed the shift operator via Eq. 9.7. Now we go the opposite direction and start out from the shift operator and derive the canonical momentum from it. This will show the intricate relationship between momenta and symmetry.

If the shift operator S for an operator $\hat{Q}$ obeying Eq. 9.8 is known, then we can construct the **canonical momentum** using

$$\begin{aligned}\hat{S}(\Delta) &= \mathrm{e}^{-\frac{i}{\hbar}\hat{P}\Delta} = \hat{1} - \frac{i}{\hbar}\hat{P}\Delta + O(\Delta^2)\\ i\hbar\frac{\hat{S}(\Delta)-\hat{1}}{\Delta} &= \hat{P} + O(\Delta)\end{aligned}$$

Thus we obtain an expression that allows us to obtain the canonical momentum to any variable, for which the corresponding shift operator is known.

CANONICAL MOMENTUM

$$\hat{P} = i\hbar \lim_{\Delta\to 0}\frac{\hat{S}(\Delta)-\hat{1}}{\Delta} \tag{9.11}$$

9.4 Recommended exercises

1. Construct the translation operator from the canonical momentum Exercise 16.4 on p. 244

Chapter 10

Specific Symmetries

10.1 Parity

An example of a finite group is the inversion:

$$
\begin{aligned}
\hat{S} &= \int dx |-x\rangle\langle x| \\
\hat{S}\psi(x) = \langle x|\hat{S}|\psi\rangle &= \langle x| \underbrace{\int dx' |-x'\rangle\langle x'|}_{\hat{S}} |\psi\rangle = \int dx' \underbrace{\langle x|-x'\rangle}_{\delta(x+x')} \underbrace{\langle x'|\psi\rangle}_{\psi(x')} \\
&= \int dx' \delta(x+x')\psi(x') = \psi(-x)
\end{aligned}
$$

The group has the elements $(\hat{1}, \hat{S})$

Since $\hat{S}^2 = \hat{1}$ we find two eigenvalues $s_1 = 1$ and $s_2 = -1$. If the eigenvalue is $+1$ we say the state has **parity** one or the state is even. If it is -1 we say the state has **parity minus one** or that the state is odd.

Example: We expect that a system is inversion symmetric, if the potential is inversion symmetric, such as the particle in the box. If we place the zero in the center of the box, then the inversion does not change the potential. As a consequence of the symmetry, the eigenstates shown in Fig. 4.4 on p. 62, are either symmetric or antisymmetric with respect to the inversion: The symmetric states have parity $s = 1$, while the antisymmetric states have parity $s = -1$.

In the following we show that the inversion symmetry can be traced back to the inversion symmetry of the potential. It seems obvious that an inversion symmetric potential leads to a inversion symmetric problem. Nevertheless this nearly trivial exercise may be helpful to demonstrate the principle of showing the symmetry of a quantum mechanical problem.

$$
\begin{aligned}
\hat{H} &= \int dx |x\rangle \left(\frac{-\hbar^2}{2m}\partial_x^2 + v(x)\right) \langle x| = \hat{T} + \hat{V} \\
\hat{T} &= \int dx |x\rangle \frac{-\hbar^2}{2m}\partial_x^2 \langle x| \\
\hat{V} &= \int dx |x\rangle v(x) \langle x|
\end{aligned}
$$

In order to verify that the system is symmetric with respect to inversion we need to show that the Hamilton operator commutes with the inversion operator.

$$
\begin{aligned}
[\hat{T},\hat{S}] &= \underbrace{\int dx|x\rangle\frac{-\hbar^2}{2m}\partial_x^2\langle x|}_{\hat{T}} \underbrace{\int dx'|-x'\rangle\langle x'|}_{\hat{S}} - \underbrace{\int dx'\;|x'\rangle\langle -x'|}_{\hat{S}} \underbrace{\int dx\;|x\rangle\frac{-\hbar^2}{2m}\partial_x^2\langle x|}_{\hat{T}} \\
&= \int dx|x\rangle\frac{-\hbar^2}{2m}\partial_x^2\int dx'\;\underbrace{\langle x|-x'\rangle}_{\delta_{x+x'}}\langle x'| - \int dx'\int dx\;|x'\rangle\underbrace{\langle -x'|x\rangle}_{\delta(x+x')}\frac{-\hbar^2}{2m}\partial_x^2\langle x| \\
&= \int dx|x\rangle\frac{-\hbar^2}{2m}\partial_x^2\langle -x| - \int dx\;|-x\rangle\frac{-\hbar^2}{2m}\partial_x^2\langle x| = 0
\end{aligned}
$$

In the last step we performed a variable transform $x \to -x$ on the right-hand part.[1] [2] Thus we have shown that the kinetic energy commutates with the inversion operator.

Now we need to investigate the potential operator

$$
\begin{aligned}
[\hat{V},\hat{S}] &= \underbrace{\int dx|x\rangle v(x)\langle x|}_{\hat{V}} \underbrace{\int dx'|-x'\rangle\langle x'|}_{\hat{S}} - \underbrace{\int dx'\;|x'\rangle\langle -x'|}_{\hat{S}} \underbrace{\int dx\;|x\rangle v(x)\langle x|}_{\hat{V}} \\
&= \int dx\int dx'|x\rangle v(x)\underbrace{\langle x|-x'\rangle}_{\delta(x+x')}\langle x'| - \int dx'\int dx\;|x'\rangle\underbrace{\langle -x'|x\rangle}_{\delta(x+x')}v(x)\langle x| \\
&= \int dx|x\rangle v(x)\langle -x| - \int dx'|x'\rangle v(-x')\langle -x'| \\
&= \int dx|x\rangle\Big(v(x)-v(-x)\Big)\langle -x|
\end{aligned}
$$

Thus the commutator $[\hat{H},\hat{S}]$ between the hamiltonian and the parity operator vanishes, if the potential is itself inversion symmetric, i.e. $v(x) = v(-x)$. Thus the Hamiltonian is symmetric under the parity operation, if the parity operation transforms the potential on itself. Remember that this is not true for the wave function: The wave functions of a parity symmetric system are either symmetric or antisymmetric under the parity operation.

10.2 n-fold rotation about an axis

Consider an n-fold rotation in space, which is described by a unitary matrix $\boldsymbol{U}$

$$\vec{r'} = \boldsymbol{U}\vec{r}$$

Since $\boldsymbol{U}$ describes a n-fold rotation, $\boldsymbol{U}^n = \mathbf{1}$.

[1] Variable transform:

$$
\begin{aligned}
y(x) = -x \qquad &\Rightarrow \qquad dy = \frac{dy}{dx}dx \overset{y=-x}{=} -dx \\
\int_a^b dx f(x) = \int_{y(a)}^{y(b)} dy\;\left(\frac{dy}{dx}\right) f(x(y)) &\overset{y=-x}{=} -\int_{-a}^{-b} dy\; f(-y) = +\int_{-b}^{-a} dy\; f(-y) \\
\partial_x^2 = \partial_x\left(\frac{dy}{dx}\partial_y\right) &= \frac{d^2y}{dx^2}\partial_y + \left(\frac{dy}{dx}\right)^2\partial_y^2 \overset{y=-x}{=} \partial_y^2
\end{aligned}
$$

[2] Here we show the identity for an arbitrary matrix element. which proves the operator identity.

$$
\begin{aligned}
\langle\phi|[\hat{T},\hat{S}]|\psi\rangle &= \int dx\langle\psi|x\rangle\frac{-\hbar^2}{2m}\partial_x^2\langle -x|\phi\rangle - \int dx\;\langle\psi|-x\rangle\frac{-\hbar^2}{2m}\partial_x^2\langle x|\phi\rangle \\
&= \int dx\;\psi^*(x)\frac{-\hbar^2}{2m}\partial_x^2\phi(-x) - \int dx\;\psi^*(-x)\frac{-\hbar^2}{2m}\partial_x^2\phi(x) = 0
\end{aligned}
$$

The rotation in real space corresponds to a transformation $\hat{S}$ in Hilbert space.

$$
\begin{aligned}
\hat{S} &= \int d^3r \, |\boldsymbol{U}\vec{r}\rangle\langle\vec{r}| \\
\hat{S}\psi(\vec{r}) &= \langle\vec{r}|\hat{S}|\psi\rangle = \int d^3r' \, \underbrace{\langle\vec{r}|\boldsymbol{U}\vec{r}'\rangle}_{\delta(\vec{r}-\boldsymbol{U}\vec{r}')} \underbrace{\langle\vec{r}'|\psi\rangle}_{\psi(\vec{r}')} \\
&\stackrel{\det[\boldsymbol{U}]=1;\boldsymbol{U}^{-1}=\boldsymbol{U}^\dagger}{=} \psi(\boldsymbol{U}^\dagger\vec{r})
\end{aligned}
$$

Note that the resolution of the delta function was non-trivial.[3]

Just as the rotation of real space point form a group with elements $\{\mathbf{1}, \boldsymbol{U}, \boldsymbol{U}^2, \ldots \boldsymbol{U}^{n-1}\}$ the group of the corresponding symmetry operations in real space forms a group. The group has n elements, $\{\hat{1}, \hat{S}, \ldots, \hat{S}^{n-1}\}$. The eigenvalues of $\hat{S}$ are $s_j = e^{i2\pi j/n}$ for $j = 0, \ldots, n-1$.

Editor: introduce here the example of ammonia, to show how eigenstates for a symmetry operator are constructed

10.3 Exchange of two particles

If the exchange of two particles does not change the system, we call those particles identical. Thus the presence of identical particles implies a symmetry. Unlike classical mechanics, this symmetry has radical consequences in quantum mechanics.

The symmetry operation is the permutation $\hat{P}$ of two particles.

$$
\hat{P}_{i,j} = \int d^3r_1 \cdots \int d^3r_N \, |\vec{r}_1, \ldots, \vec{r}_j, \ldots, \vec{r}_i, \ldots, \vec{r}_N\rangle\langle\vec{r}_1, \ldots, \vec{r}_i, \ldots, \vec{r}_j, \ldots, \vec{r}_N|
$$

The wave function that corresponds to the state $|\vec{r}_1, \ldots, |\vec{r}_N\rangle$ is a product of δ-functions in the 3N-dimensional configuration space centered at the positions $\vec{r}_1, \ldots, \vec{r}_N$, i.e.

$$
|\psi\rangle = |\vec{r}_1, \ldots, \vec{r}_N\rangle \qquad \Rightarrow \qquad \psi(\vec{r'}_1, \ldots, \vec{r'}_N) = \langle\vec{r'}_1, \ldots, \vec{r'}_N|\vec{r}_1, \ldots, \vec{r}_N\rangle = \delta(\vec{r'}_1 - \vec{r}_1)\cdots\delta(\vec{r'}_N - \vec{r}_N)
$$

The permutation can be considered as a geometric operation in the $3N$-dimensional configuration space of all N particles. This is illustrated best by considering two one-dimensional particles as the simplest possible case. The positions of the two particles are described by two coordinates (x_1, x_2). The permutation of the two particles corresponds to a mirror operation about the main diagonal of the $x_1 - x_2$-plane. This is illustrated in Fig. 10.3.1 on p. 152.

Let us now consider a wave function $\psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$ for many particles and the permutation of two particles as symmetry operator.

$$
\hat{P}_{i,j}\psi(\ldots, \vec{r}_i, \ldots, \vec{r}_j \ldots) = \psi(\ldots, \vec{r}_j, \ldots, \vec{r}_i \ldots)
$$

For the interchange of every two particles we have a symmetry group with two elements and eigenvalues 1 and -1.

For identical particles there exists no Hamilton operator that is not invariant with respect to particle exchange. If there would be one, then it would be possible to devise and experiment that distinguishes between the particles. If we can distinguish between particles, they are obviously no more be identical. Since this is not the case, each Hamilton operator commutates with the particle exchange of identical particles. Thus all particles are either symmetric or antisymmetric with respect to particle exchange. Accordingly, particles can be classified as being either Fermions or Bosons.

[3] In one dimension the identity $\delta(ax) = \frac{1}{a}\delta(x)$ holds. In three dimension this translates into $\delta(\boldsymbol{A}\vec{x}) = \frac{1}{\det \boldsymbol{A}}\delta(\vec{x})$. In the operation used in the text, the prefactor vanishes, because $\boldsymbol{U}$ is unitary and therefore $\det \boldsymbol{U} = 1$

FERMIONS AND BOSONS

Fermions	Bosons
wave function antisymmetric under permutation	wave function symmetric under permutation
half-integer spin	integer spin
Pauli principle	macroscopic quantum states possible
	(e.g. Bose Einstein condensation, Laser)
"particles"	"interactions"
Examples:	Examples
electron	photon (electrodynamics)
proton	graviton (gravity)
neutron	Z- and W-boson (weak force)
quark	gluon (strong force)
	Higgs boson
	phonon
	exciton

10.3.1 Fermions

Particles that are antisymmetric under particle exchange are called **Fermions**. All Fermions have a half-integer spin. (The concept of spin will be introduced later.) Examples for Fermions are electrons, protons, quarks. An important consequence of the antisymmetry is the Pauli principle, which says that no two Fermions can occupy the same site or the same one-particle state.

Let us show that two identical Fermions cannot occupy the same site. The wave function amplitude for two particles at the same position $\vec{r}_0$ and all others on arbitrary positions is

$$\psi(\ldots,\underbrace{\vec{r}^{(0)}}_{pos.i},\ldots,\underbrace{\vec{r}^{(0)}}_{pos.j},\ldots)$$

Exchanging the two particle coordinates, one obtains , on the one hand, the same result, on the other hand we obtain a sign change due to particle permutation. Thus:

$$\psi(\ldots,\underbrace{\vec{r}^{(0)}}_{pos.i},\ldots,\underbrace{\vec{r}^{(0)}}_{pos.j},\ldots)=-\psi(\ldots,\underbrace{\vec{r}^{(0)}}_{pos.i},\ldots,\underbrace{\vec{r}^{(0)}}_{pos.j},\ldots)$$

This is only possible if the wave function amplitude vanishes, whenever two particles have the same position.[4] The probability that the two particles are at $\vec{r}_0$ is obtained from the absolute square of the wave function integrated over all particle positions except particles i and j. This probability vanishes as well.

PAULI PRINCIPLE

- The probability of finding two identical Fermions in the same position vanishes
- Two identical Fermions cannot occupy the same one-particle state.

Let us now consider a product state $\Psi(\vec{r}_1,\vec{r}_2)=\phi_1(\vec{r}_1)\phi_2(\vec{r}_2)$. Product states are the most simple many particle wave functions. They play an important role for non-interacting particles.

[4]If the two particles are at the same position the value of the wave function does not change under permutation. On the other hand the wave function has to change its sign. Thus the wave function vanishes for identical positions.

For non-interacting particles the N-particle Hamiltonian $\hat{H}^{(N)} = \sum_{i=1}^{N} \hat{H}_i^{(1)}$ is a sum of one-particle Hamiltonians, where the one-particle Hamiltonian $\hat{H}_i^{(1)}$ only acts on the coordinates of the i-th particle. For non-interacting particles, we can use the method of separation of variables to express the eigenstates as a product of wave functions of the one-particle wave functions, that is a product state. The energy is then the sum of the corresponding one-particle energies.

However, a product state does not have the symmetry or antisymmetry required for identical particles. Thus we need to antisymmetrize the wave function. We can anti-symmetrize this state an according to our recipe from Eq. 9.6. In the two-particle case we obtain

$$\Psi^F(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2!}}\Big[\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) - \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)\Big]$$

Here we can easily verify that the Pauli principle holds: the wave function vanishes if $\phi_1(r) = \phi_2(r)$. Thus we have shown that one cannot place two identical Fermions into a single one-particle state.

An antisymmetric wave function can be written in a more general way as a **Slater determinant**

$$\Psi^F(\vec{r}_1, \dots, \vec{r}_N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \phi_1(\vec{r}_1) & \phi_1(\vec{r}_2) & \cdots & \phi_1(\vec{r}_N) \\ \phi_2(\vec{r}_1) & \phi_2(\vec{r}_2) & \cdots & \phi_2(\vec{r}_N) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_N(\vec{r}_1) & \phi_N(\vec{r}_2) & \cdots & \phi_N(\vec{r}_N) \end{pmatrix}$$

From the property of that the determinant it changes its sign under exchange of two columns, we obtain antisymmetry. Not every antisymmetric wave function can be represented as a Slater determinant. The most general form of a Fermionic wave function of N particles is a superposition of all N-particle Slater determinants that can be formed from a complete set of one-particle states.

Not every N-particle state can be expressed as a single Slater determinant. However, if we form all possible the Slater determinants from the eigenstates of a one-particle Hamiltonian $\hat{h}$, we obtain a complete basisset of the Hilbert space of Fermionic, i.e. antisymmetric, N-particle states. These Slater-determinants are the eigenstates of the non-interacting N-particle Hamiltonian $\hat{H}^{(N)} = \sum_{i=1}^{N} \hat{H}_i^{(1)}$. For identical particles, each one-particle Hamiltonian $\hat{H}_i^{(1)}$ is identical to all others except that it acts on the coordinates of a specified, namely the i-th, particle. The eigenstates of an observable such as the N-particle Hamiltonian forms a complete basisset. A general N-particle wave function can thus be written as a **Multiconfiguration wave function**, namely as a sum of all Slater determinants that can be built from a complete one-particle basisset.

The energy eigenstates of non-interacting Fermions can be expressed as single Slater determinants made from the eigenstates of the corresponding one-particle Hamiltonian. The wave functions of interacting particles must be written as multiconfiguration wave functions. Multiconfiguration wave functions are notoriously difficult to describe. However there are techniques such as density functional theory, that allow to map a system of interacting electrons onto a system of non-interacting particles. The latter allow us to effectively deal with wave functions that are single Slater determinants.

10.3.2 Bosons

Particles that are symmetric under particle exchange are called **Bosons**. All Bosons have an integer spin. (The concept of spin will be introduced later.) Examples for bosons are photons, mesons. Typically they are related to interactions such as the Coulomb interaction.

I can form a Bosonic state as a symmetrized product of a single one-particle states. For a two particle state built from two one-particle orbitals $\phi_1(\vec{r})$ and $\phi_2(\vec{r})$one obtains

$$\Psi^B(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}}\left(\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) + \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)\right)$$

This can be generalized to many particles as follows:

$$\Psi^B(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \prod_{i<j} \frac{1}{\sqrt{2}}(1 + P_{i,j})\phi(\vec{r}_1)\phi(\vec{r}_2)\dots\phi(\vec{r}_N)$$

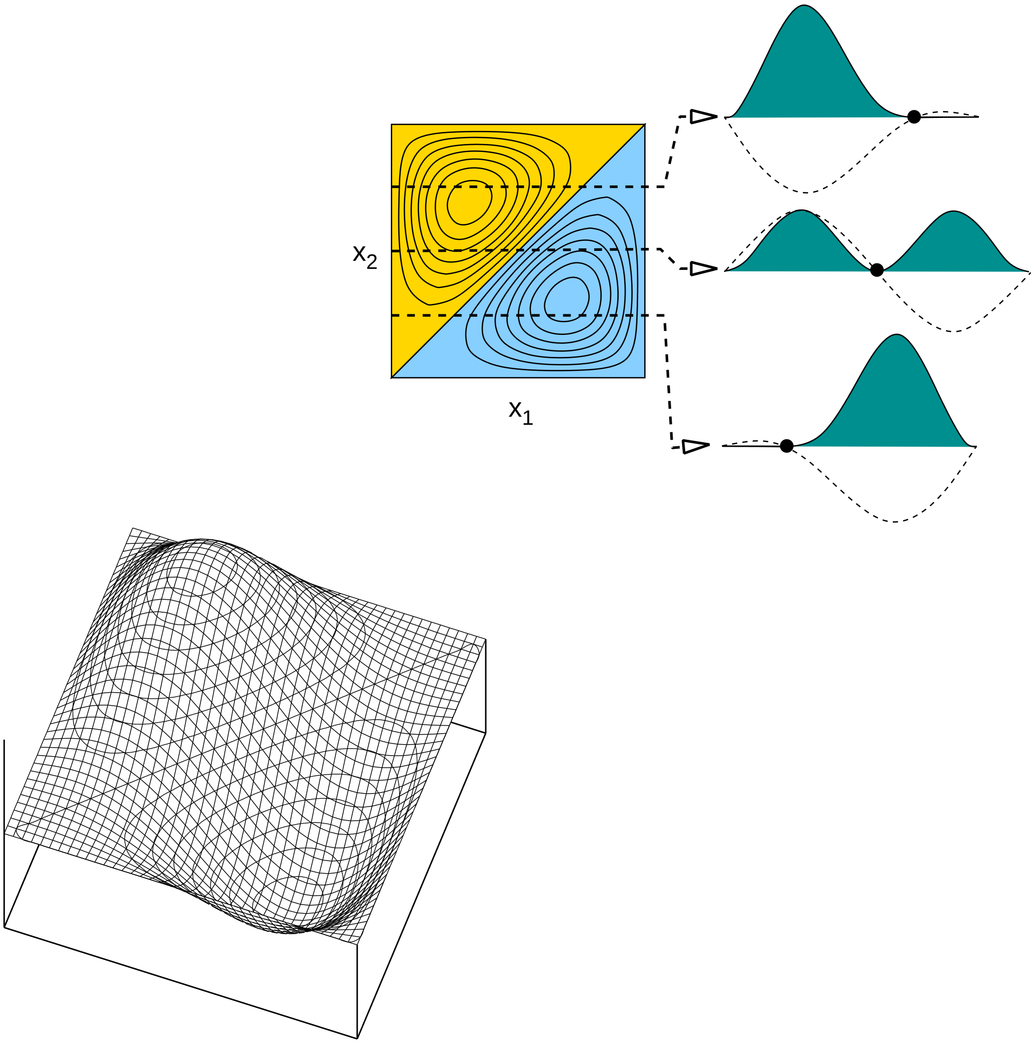


Fig. 10.1: Wave function of two fermions in a one-dimensional box. One axis represents the position of the first particle and the other axis that of the other. The problem of two particles in a one-dimensional box is analogous to that of one-particle in a two dimensional box. The symmetry with respect to particle exchange corresponds to a mirror plane at the diagonal $x_1 = x_2$. The antisymmetry forces a node plane to lie at the diagonal $x_1 = x_2$. The lower figure shows a contour plot of the wave function. The different signs are indicated by the different colors. The functions on the right indicate the wave function (dashed) and the density (filled line) of the first particle for a fixed position of the second. The position of the second is marked by the black ball.

where the operator $P_{i,j}$ permutes the coordinates $\vec{r}_i$ and $\vec{r}_j$ Thus we can place an macroscopic number of Bosons into a single one-particle state. Such a state is called a **macroscopic quantum** state, and it allows quantum properties to be observed in the macroscopic . An example for such a state is laser light. Another example is the superfluid state of liquid Helium, which flows without dissipation and has infinite viscosity. Superconductivity is the result of conduction electrons forming pairs, which are Bosons, and the condensation of these pairs into a single one-particle orbital. Even Atoms can be prepared in a Bosonic state, which undergoes Bose-Einstein condensation and then all atoms are placed in a macroscopic quantum state.

10.4 Continuous translations

The translation operator, which translates a state in x-direction, is defined by

$$\hat{S} = \int dx\, |x+\Delta\rangle\langle x|$$
$$\hat{S}(\Delta)\Psi(x) = \langle x|\hat{S}|\psi\rangle = \int dx'\, \underbrace{\langle x|x'+\Delta\rangle}_{\delta(x'+\Delta-x)}\langle x'|\psi\rangle = \langle x-\Delta|\psi\rangle = \Psi(x-\Delta)$$

Note here the minus sign in the argument: If $\hat{S}(\Delta)$ shifts the wave function to the right, the values for the new wave function are obtained from the unshifted wave function at a position to the left.

We obtain the conjugate momentum by

$$\begin{aligned}
\hat{p}_x &\stackrel{\text{Eq. 9.11}}{=} i\hbar \lim_{\Delta\to 0}\frac{1}{\Delta}\int dx\, \Big(|x+\Delta\rangle\langle x| - |x\rangle\langle x|\Big) \\
&= i\hbar \lim_{\Delta\to 0}\frac{1}{\Delta}\int dx\, \Big(|x\rangle\langle x-\Delta| - |x\rangle\langle x|\Big) \\
&= i\hbar \int dx |x\rangle \underbrace{\lim_{\Delta\to 0}\frac{\langle x-\Delta| - \langle x|}{\Delta}}_{-\partial_x\langle x|} \\
&= \int dx\, |x\rangle\frac{\hbar}{i}\partial_x\langle x| \\
\hat{p}_x\Psi(x) &= \langle x|\hat{p}_x|\psi\rangle = \int dx'\, \underbrace{\langle x|x'\rangle}_{\delta(x-x')}\frac{\hbar}{i}\partial_{x'}\underbrace{\langle x'|\psi\rangle}_{\psi(x')} \\
&= \frac{\hbar}{i}\partial_x\psi(x)
\end{aligned}$$

Thus we have derived the explicit form for the momentum operator from the shift operator.

The generalization to a translation in three dimensions is straightforward:

$$\hat{S}(\vec{\Delta}) = \int d^3r\, |\vec{r}+\vec{\Delta}\rangle\langle\vec{r}|$$
$$\hat{S}(\vec{\Delta})\Psi(\vec{r}) = \Psi(\vec{r}-\vec{\Delta})$$
$$\hat{\vec{p}} = \int d^3r\, |\vec{r}\rangle\frac{\hbar}{i}\vec{\nabla}\langle\vec{r}|$$

The eigenstates $|\psi_{\vec{k}}\rangle$ of the shift operator can be obtained from the eigenvalue equation of the momentum operator

$$\hat{p}\psi_{\vec{k}}(\vec{r}) = \psi_{\vec{k}}(\vec{r})\hbar\vec{k} \qquad \Rightarrow \qquad \hat{S}(\vec{\Delta})\Psi_{\vec{k}}(\vec{r}) = \Psi_{\vec{k}}(\vec{r}-\vec{\Delta}) \stackrel{\text{Eq. 9.7}}{=} \Psi_{\vec{k}}(\vec{r})\mathrm{e}^{-i\vec{k}\vec{\Delta}} \tag{10.1}$$

where $\hbar\vec{k}$ are here the eigenvalues of the momentum operator. By setting $\vec{r}=\vec{0}$ and replacing $\vec{\Delta}$ by $-\vec{r}$ in Eq. 10.1, we find that the eigenstates of the translation operator, which are also momentum eigenstates have the from of plane waves.

$$
\begin{aligned}
& & \Psi_{\vec{k}}(\vec{r}-\vec{\Delta}) &\stackrel{\text{Eq. 10.1}}{=} \Psi_{\vec{k}}(\vec{r})\mathrm{e}^{-i\vec{k}\vec{\Delta}} \\
\stackrel{\vec{r}\to\vec{0};\ \vec{\Delta}\to-\vec{r}}{\Rightarrow} & & \Psi_{\vec{k}}(\vec{0}-(-\vec{r})) &= \Psi_{\vec{k}}(\vec{0})\mathrm{e}^{-i\vec{k}(-\vec{r})} \\
& \Rightarrow & \Psi_{\vec{k}}(\vec{r}) &= \Psi_{\vec{k}}(\vec{0})\mathrm{e}^{i\vec{k}\vec{r}}
\end{aligned}
\tag{10.2}
$$

EIGENSTATES OF THE TRANSLATION OPERATOR

The eigenstates of the translation operator,

$$\hat{S}(\vec{\Delta})|\psi_{\vec{k}}\rangle = |\psi_{\vec{k}}\rangle \mathrm{e}^{-i\vec{k}\vec{\Delta}} \tag{10.3}$$

are plane waves

$$\Psi_{\vec{k}}(\vec{r}) \stackrel{\text{Eq. 10.2}}{=} A\,\mathrm{e}^{i\vec{k}\vec{r}} \tag{10.4}$$

The eigenstates of the shift operator are also the eigenstates of the corresponding momentum operator, i.e.

$$\hat{p}|\psi_{\vec{k}}\rangle = |\psi_{\vec{k}}\rangle\hbar\vec{k}$$

This is the reason for using plane waves as Ansatz, when we search a solution for a differential equation with translational symmetry. Because the result rests on symmetry, this Ansatz is very general and works for any kind of differential equation.

10.5 Discrete translations, crystals and Bloch theorem

Most solids are crystals. This implies that the atoms are located on regular positions. The clearest evidence of this regular arrangement are the crystal faces of gem stones. This arrangement results in a symmetry, the lattice periodicity. Lattice translation symmetry is a discrete translational symmetry.

10.5.1 Real and reciprocal lattice: One-dimensional example

Let us consider a linear chain of hydrogen atoms as the most simple model of a one-dimensionally periodic crystal of hydrogen atoms shown in Fig. 10.2. It is a model system, which is not stable in reality. However it allows to develop the concepts.

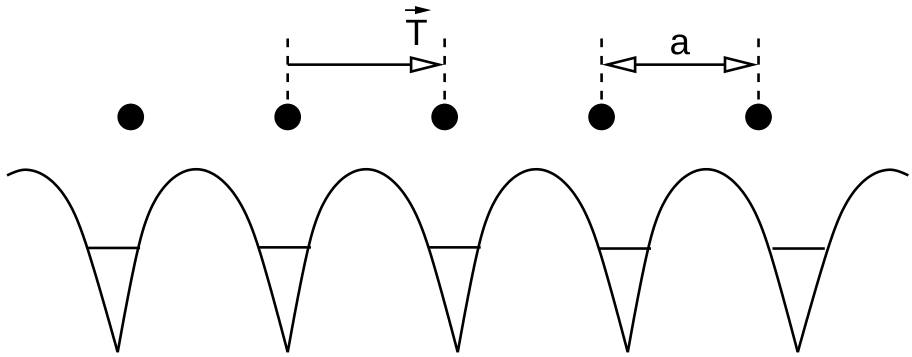


Fig. 10.2: Linear chain of hydrogen atoms. The lattice spacing is a. The balls represent the atoms. The graph sketched below represents the potential as function of position. The primitive lattice vector is $\vec{T}$ which has the length a. The primitive reciprocal lattice vector has the length $2\pi/a$.

The periodicity of the chain implies that for every atom at position $\vec{R}_0$ there are also equivalent atoms at the positions $\vec{R}_n = \vec{R}_0 + n\vec{T}$, where n is any integer. $\vec{T}$ is the **primitive lattice vector**. **General lattice vectors** are $\vec{t}_n = n\vec{T}$. The primitive lattice vector is the smallest possible lattice vector, that fully describes the lattice periodicity. Its length is called the **lattice constant** a.

Not only the atomic positions are regular, but also the potential is **periodic**, that is $v(\vec{r}+\vec{t}_n) = v(\vec{r})$. A periodic potential has a discrete Fourier transform

$$v(\vec{r}) = \sum_n \mathrm{e}^{i\vec{G}_n\vec{r}} V_n$$

where the vectors $\vec{G}_n = n\vec{g}$ are the **general reciprocal lattice vectors**. The **primitive reciprocal lattice vector** is denoted by $\vec{g}$. The length of the primitive reciprocal lattice vector is inversely proportional to the real-space lattice constant a, that is $|\vec{g}| = \frac{2\pi}{a}$

10.5.2 Real and reciprocal lattice in three dimensions

Let us now generalize the expressions to three a three-dimensional lattice. Fig. 10.5.2 demonstrates the concepts developed in the following. Corresponding to the three spatial directions there are now three primitive lattice vectors, which we denote by $\vec{T}_1, \vec{T}_2, \vec{T}_3$. Note that in two or three dimensions there is no unique choice of primitive lattice vectors. The three primitive lattice vectors span the **primitive unit cell**.

A general lattice vector $\vec{t}_{i,j,k}$ can be expressed by the **primitive lattice vectors** $\vec{T}_1, \vec{T}_2, \vec{T}_3$ as

$$\vec{t}_{i,j,k} = i\vec{T}_1 + j\vec{T}_2 + k\vec{T}_3 \hat{=} \underbrace{\begin{pmatrix} T_{x,1} & T_{x,2} & T_{x,3} \\ T_{y,1} & T_{y,2} & T_{y,3} \\ T_{z,1} & T_{z,2} & T_{z,3} \end{pmatrix}}_{\boldsymbol{T}} \begin{pmatrix} i \\ j \\ k \end{pmatrix}$$

where i, j, k are arbitrary integers.

It is often convenient to combine the lattice vectors into a 3×3 matrix $\boldsymbol{T}$ as shown above. Often the atomic positions $\vec{R}_n$ are provided in **relative positions** $\vec{X}$ which are defined as

$$\vec{R}_n = \boldsymbol{T}\vec{X}_n \qquad \Leftrightarrow \qquad \vec{X}_n = \boldsymbol{T}^{-1}\vec{R}_n$$

A potential is called periodic with respect to these lattice translations, if $V(\vec{r}+\vec{t}_{i,j,k}) = V(\vec{r})$.

RECIPROCAL LATTICE

The reciprocal lattice is given by those values of the wave vector $\vec{G}$, for which the corresponding plane waves $e^{i\vec{G}\vec{r}}$ have the same periodicity as the real-space lattice.
The primitive reciprocal-lattice vectors $\vec{g}_n$ for $n = 1, 2, 3$ are defined in three dimensions as

$$\vec{g}_n \vec{T}_m = 2\pi\delta_{n,m} \tag{10.5}$$

Thus, in three dimensions the reciprocal lattice vectors can be obtained as

$$\vec{g}_1 = 2\pi\frac{\vec{T}_2 \times \vec{T}_3}{\vec{T}_1\left(\vec{T}_2 \times \vec{T}_3\right)} \quad , \quad \vec{g}_2 = 2\pi\frac{\vec{T}_3 \times \vec{T}_1}{\vec{T}_1\left(\vec{T}_2 \times \vec{T}_3\right)} \quad , \quad \vec{g}_3 = 2\pi\frac{\vec{T}_1 \times \vec{T}_2}{\vec{T}_1\left(\vec{T}_2 \times \vec{T}_3\right)}$$

It is easily shown that these expressions fulfill the defining equation Eq. 10.5 for the reciprocal lattice vectors. Note that the expressions for the second and third lattice vector are obtained from the first by cyclic commutation of the indices.

Because the vector product is perpendicular to the two vectors forming it, every reciprocal lattice vector is perpendicular to two real-space lattice vectors. This brings us to the physical meaning of the reciprocal lattice vectors: Two real space lattice vectors define a lattice plane. A plane can be defined either by two linearly independent vectors in the plane or alternatively by the plane normal. The reciprocal lattice vectors apparently define lattice planes, because they are plane normals.

Let us now consider the distance Δ_n of two neighboring lattice planes. It is obtained by projecting one vector pointing $\vec{T}_n$ from one plane to the other onto a normalized plane normal $\vec{g}_n/|\vec{g}_n|$.

$$\begin{aligned} \Delta_n &= \vec{T}_n\frac{\vec{g}_n}{|\vec{g}_n|} \stackrel{\text{Eq. 10.5}}{=} \frac{2\pi}{|\vec{g}_n|} \\ \Rightarrow \qquad |\vec{g}_n| &= \frac{2\pi}{\Delta_n} \end{aligned}$$

PHYSICAL MEANING OF RECIPROCAL LATTICE VECTORS

- The reciprocal lattice vectors are perpendicular to the lattice planes of the real-space lattice.
- The length of the *primitive* reciprocal lattice vectors is 2π divided by the distance of the lattice planes.

We can form a 3×3 matrix $\boldsymbol{g}$ from the three primitive reciprocal lattice vectors.

$$\boldsymbol{g} \stackrel{\text{def}}{=} \begin{pmatrix} g_{x,1} & g_{x,2} & g_{x,3} \\ g_{y,1} & g_{y,2} & g_{y,3} \\ g_{z,1} & g_{z,2} & g_{z,3} \end{pmatrix}$$

The definition Eq. 10.5 of the reciprocal lattice vectors can be expressed in matrix form as

$$\boldsymbol{g}^{\top}\boldsymbol{T} \stackrel{\text{Eq. 10.5}}{=} 2\pi\mathbf{1} \tag{10.6}$$

Thus the matrix $\boldsymbol{g}$ is obtained as

$$\boldsymbol{g} \stackrel{\text{Eq. 10.6}}{=} 2\pi\left(\boldsymbol{T}^{-1}\right)^{\top} \tag{10.7}$$

The general reciprocal lattice vectors are

$$\vec{G}_{i,j,k} = i\vec{g}_1 + j\vec{g}_2 + k\vec{g}_3$$

The reciprocal lattice vectors play an important role in the Fourier transform of periodic functions. All plane waves that are periodic with the lattice vectors $\vec{T}_n$ are characterized by a reciprocal lattice vector.

$$\begin{aligned}
\mathrm{e}^{i\vec{G}(\vec{r}+\vec{T}_n)} &= \mathrm{e}^{i\vec{G}\vec{r}} \\
\Rightarrow \vec{G}(\vec{r}+\vec{T}_n) &= \vec{G}\vec{r} + 2\pi m_n \\
\Rightarrow \vec{G}\vec{T}_n &= 2\pi m_n \\
\Rightarrow \vec{G}\boldsymbol{T} &= 2\pi \vec{m} \\
\Rightarrow \vec{G} &= 2\pi \vec{m}\boldsymbol{T}^{-1} = 2\pi \left(\boldsymbol{T}^{-1}\right)^{\top} \vec{m}
\end{aligned}$$

where m_n are integers. The allowed values for $\vec{G}$ are exactly those of the reciprocal lattice. This was the motivation for the definition Eq. 10.6.

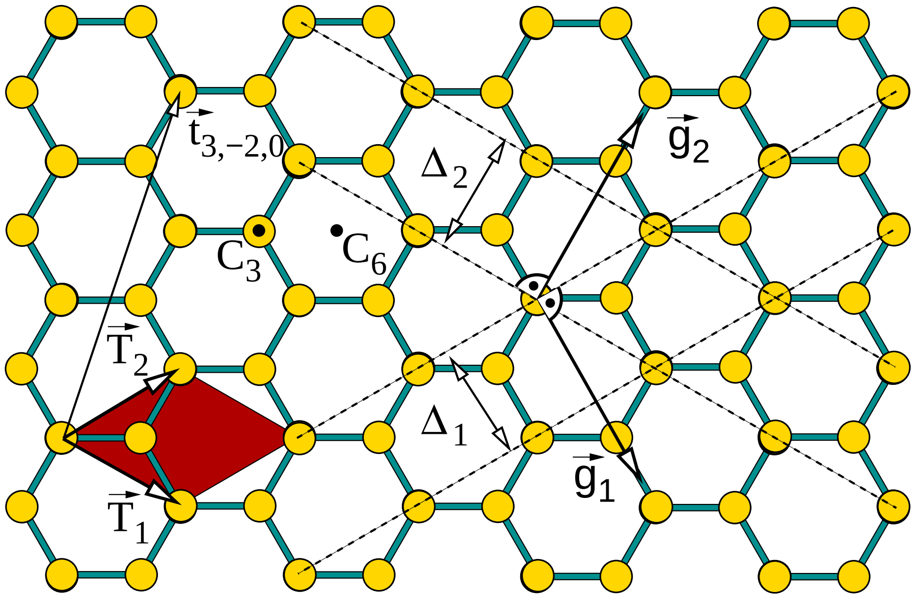


Fig. 10.3: Translational and rotational symmetry of a single graphite sheet and demonstration of reciprocal lattice vectors. Graphite is a layered material, which consists of sheets as the one shown. The yellow balls indicate the carbon atoms and the blue sticks represent the bonds. In graphite these two-dimensional sheets are stacked on top of each other, where every second sheet is shifted such that an atom falls vertically below the point indicated by C_6. Here we only consider the symmetry of a single sheet. The elementary unit cell is indicated by the red parallelogram, which is spanned by the elementary lattice vectors $\vec{T}_1$ and $\vec{T}_2$. The third lattice vector points perpendicular to the sheet towards you. An example for a general lattice vector is $\vec{t}_{3,-2,0} = 3\vec{T}_1 - 2\vec{T}_2 + 0\vec{T}_3$. The lattice planes indicated by the dashed lines. The lattice planes are perpendicular to the sheet. The distance of the lattice planes are indicated by Δ_1 and Δ_2. The elementary reciprocal lattice vectors are $\vec{g}_1$ and $\vec{g}_2$. The third reciprocal lattice vector points perpendicular out of the plane towards you. Note that the reciprocal lattice vectors have the unit "inverse length". Thus their length is considered irrelevant in this real space figure. The axis through C_3 standing perpendicular to the plane is a three-fold rotation axis of the graphite crystal. The axis through C_6 perpendicular to the plane is a 6-fold rotation axis of the graphite sheet, but only a three-fold rotation axis of the graphite crystal. (Note that a rotation axis for the graphite crystal must be one for both sheets of the crystal). In addition there is a mirror plane lying in the plane. Furthermore there are several mirror planes perpendicular to the plane: One passing through every atom with one of the three bonds lying in the plane and one perpendicular each bond passing through the bond center.

10.5.3 Bloch Theorem

BLOCH THEOREM:

The eigenstates of a system with lattice symmetry can be written as product of a periodic function $u_k(\vec{r})$ and a modulating plane wave $e^{i\vec{k}\vec{r}}$, that is

$$\Psi_k(\vec{r}) \stackrel{\text{def}}{=} u_k(\vec{r})e^{i\vec{k}\vec{r}} \tag{10.8}$$

where the **Bloch vector** $\vec{k}$ is a good quantum number of the system.
More precisely: Lattice translation symmetry implies that there are lattice translation vectors $\vec{T}_1, \vec{T}_2, \vec{T}_3$, so that the potential obeys $V(\vec{r}+\vec{T}_i) = V(\vec{r})$. In that case the periodic function obeys $u_{\vec{k}}(\vec{r}+\vec{T}_i) = u_{\vec{k}}(\vec{r})$, and the Bloch vectors $\vec{k}$ lie in one reciprocal unit cell.

PROOF: Because the system has lattice symmetry, the Hamilton operator commutates with the lattice translation operator(Eq. 9.2). As a consequence we can form common eigenstates of the lattice translation operator and the Hamilton operator. Thus the Hamilton operator is block diagonal in the basis of eigenstates of the translation operator (Eq. 9.3). The eigenstates of the Hamilton operator can then be obtained by diagonalizing the Hamilton operator in the subspace of all eigenvalues of the translation operator with a given eigenvalue of the lattice translation vector. This is to be repeated for every eigenvalue of the lattice translation operator. All this follows directly from what we have derived already.
Now we need to find the eigenstates and eigenvalues for the lattice translation vector. The lattice translation operator is defined by

$$\hat{S}_{\vec{T}_n} = \int d^3r |\vec{r}+\vec{T}_n\rangle\langle\vec{r}|$$

where $\vec{T}_n$ for $n = 1, 2, 3$ are the three primitive lattice translation vectors in real space. The result on a wave function is

$$\hat{S}_{\vec{T}_n}\Psi(\vec{r}) \stackrel{\text{def}}{=} \langle\vec{r}|\hat{S}_{\vec{T}_n}|\Psi\rangle = \Psi(\vec{r}-\vec{T}_n)$$

We need not consider general translations, because the lattice translation operator for a general translation vector $\vec{t}_{i,j,k} = i\vec{T}_1 + j\vec{T}_2 + k\vec{T}_3$ is the product of the those for the primitive lattice translation vectors.

$$\hat{S}_{\vec{t}_{i,j,k}} = \hat{S}^i_{\vec{T}_1}\hat{S}^j_{\vec{T}_2}\hat{S}^k_{\vec{T}_3}$$

The lattice translation operator used for the lattice symmetry is the same as the continuous lattice translation vector, specialized to the lattice translations. The eigenvalues for the general lattice translation vector are already known. The eigenvalue equation is

$$\psi(\vec{r}-\vec{T}_n) = \hat{S}_{\vec{T}_n}\psi(\vec{r}) \stackrel{\text{Eq. 10.3}}{=} \psi(\vec{r})e^{-i\vec{k}\vec{T}_n}$$
$$\stackrel{\cdot e^{-i\vec{k}(\vec{r}-\vec{T}_n)}}{\Rightarrow} \psi(\vec{r}-\vec{T}_n)e^{-ik(\vec{r}-\vec{T}_n)} = \psi(\vec{r})e^{-i\vec{k}\vec{r}}$$

Thus the function

$$u_k(\vec{r}) \stackrel{\text{def}}{=} \psi(\vec{r})e^{-i\vec{k}\vec{r}} \tag{10.9}$$

is periodic, that is

$$u_k(\vec{r}-\vec{T}_n) = u_k(\vec{r})$$

This is already what had to be proven: The eigenstates of the lattice translation operator for a given eigenvalue $e^{-i\vec{k}\vec{T}_n}$, and thus the eigenstates of the Hamilton operator, have the form of a product of a periodic function $u_k(\vec{r})$ and modulated by a plane wave $e^{i\vec{k}\vec{r}}$. The wave function has the form

$$\Psi_k(\vec{r}) \stackrel{\text{Eq. 10.9}}{=} u_k(\vec{r})e^{i\vec{k}\vec{r}}$$

q.e.d

We can define a k-dependent Hamiltonian for the periodic part $u_{\vec{k}}(\vec{r})$ of the Bloch waves.

$$\begin{aligned}
\hat{H}\mathrm{e}^{i\vec{k}\vec{r}}u_{\vec{k}}(\vec{r}) &= \mathrm{e}^{i\vec{k}\vec{r}}u_{\vec{k}}(\vec{r})E_{n,\vec{k}} \\
[\mathrm{e}^{-i\vec{k}\vec{r}}\hat{H}\mathrm{e}^{i\vec{k}\vec{r}}]u_{\vec{k}}(\vec{r}) &= u_{\vec{k}}(\vec{r})E_{n,\vec{k}} \\
\hat{H}(\vec{k}) &= \mathrm{e}^{-i\vec{k}\vec{r}}\hat{H}\mathrm{e}^{i\vec{k}\vec{r}} = \mathrm{e}^{-i\vec{k}\vec{r}}[\frac{\vec{p}^2}{2m}+V(\vec{r})]\mathrm{e}^{i\vec{k}\vec{r}} = \frac{(\vec{p}+\hbar\vec{k})^2}{2m}+V(\vec{r}) \\
&= \frac{-\hbar^2}{2m}(\vec{\nabla}+i\vec{k})^2+V(\vec{r})
\end{aligned}$$

The energies $E_n(\vec{k})$ are called band structure. They are the dispersion relations of the particles in a periodic potential. Band structures exist for electrons, phonons and photons.

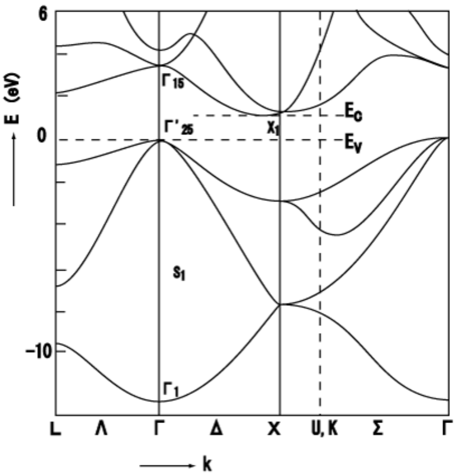


Fig. 10.4: Schematics of a calculated band structure of crystalline silicon [33].

The band structure is periodic in k-space. We show this as follows: We begin with some Bloch wave $\Psi_{\vec{k},n}(\vec{r})$, which shall be also an eigenstate of the Hamiltonian with energy $E_n(\vec{k})$. We multiply the periodic function $u_{\vec{k},n}(\vec{r})$ with $\mathrm{e}^{-i\vec{G}\vec{r}}$, which is itself periodic, and compensate by multiplication of the phase factor with $\mathrm{e}^{i\vec{G}\vec{r}}$.

$$\Psi_{\vec{k},n}(\vec{r}) = u_{\vec{k},n}(\vec{r})\mathrm{e}^{i\vec{k}\vec{r}} = [u_{\vec{k},n}(\vec{r})\mathrm{e}^{-i\vec{G}\vec{r}}]\mathrm{e}^{i(\vec{k}+\vec{G})\vec{r}}$$

Thus we have written the very same state as a Bloch wave

$$\psi_{\vec{k}+\vec{G},n}(\vec{r}) = u_{\vec{k}+\vec{G},n}\mathrm{e}^{i(\vec{k}+\vec{G})\vec{r}}$$

for the k-point $\vec{k}+\vec{G}$, with a periodic part

$$u_{\vec{k}+\vec{G},n}(\vec{r}) \stackrel{\text{def}}{=} u_{\vec{k}}(r)\mathrm{e}^{-i\vec{G}\vec{r}}$$

and a phase $\mathrm{e}^{i(\vec{k}+\vec{G})\vec{r}}$.

Because both function are identical, i.e. $\Psi_{\vec{k},n}(\vec{r}) = \Psi_{\vec{k}+\vec{G},n}(\vec{r})$, also the energies are the same. This shows that the band structure $E_n(\vec{k})$ is periodic in reciprocal space, i.e

$$E_n(\vec{k}) = E_n(\vec{k}+\vec{G})$$

10.5.4 Schrödinger equation in momentum representation

Editor: Construction! We can use the periodicity of the band structure to simplify its graphical representation. Let us consider a free particle. Clearly a completely translationally symmetric system also has lattice symmetry. The dispersion relation of a free particle is a parabola.

The wave functions of the free particle have the form

$$\phi_{\vec{G}}(\vec{r}) = \mathrm{e}^{i\vec{G}\vec{r}} = \langle\vec{r}|\vec{G}\rangle$$

and the eigenvalues are

$$E_{\vec{G}} = \frac{(\hbar\vec{G})^2}{2m}$$

Now we can consider the Hamiltonian in plane waves. A general state can be expressed in its Fourier representation $|\vec{G}\rangle$ as

$$|\Psi\rangle = \int \frac{d^3G}{(2\pi)^3}|\vec{G}\rangle c(\vec{G})$$

$$\int \frac{d^3G'}{(2\pi)^3}\langle\vec{G}|\hat{H}-E|\vec{G}'\rangle c(\vec{G}) = 0$$

$$\int \frac{d^3G'}{(2\pi)^3}\langle\vec{G}|\int d^3r\ \underbrace{|\vec{r}\rangle\left(-\frac{\hbar^2}{2m}\vec{\nabla}^2+V(\vec{r})-E\right)\langle\vec{r}|}_{\hat{H}}|\vec{G}'\rangle c(\vec{G}) = 0$$

$$\int \frac{d^3G'}{(2\pi)^3}\int d^3r\ \underbrace{\langle\vec{G}|\vec{r}\rangle}_{\mathrm{e}^{-i\vec{G}\vec{r}}}\left(-\frac{\hbar^2}{2m}\vec{\nabla}^2+V(\vec{r})-E\right)\underbrace{\langle\vec{r}|\vec{G}'\rangle}_{\mathrm{e}^{i\vec{G}'\vec{r}}} c(\vec{G}) = 0$$

$$\int \frac{d^3G'}{(2\pi)^3}\int d^3r\ \mathrm{e}^{-i\vec{G}\vec{r}}\left(-\frac{\hbar^2}{2m}\vec{\nabla}^2+V(\vec{r})-E\right)\mathrm{e}^{i\vec{G}'\vec{r}}c(\vec{G}) = 0$$

$$\int \frac{d^3G'}{(2\pi)^3}\int d^3r\ \left(\frac{\hbar^2\vec{G}'^2}{2m}+V(\vec{r})-E\right)\mathrm{e}^{i(\vec{G}'-\vec{G})\vec{r}}c(\vec{G}) = 0$$

$$\int \frac{d^3G'}{(2\pi)^3}\left(\frac{\hbar^2\vec{G}'^2}{2m}(2\pi)^3\delta(\vec{G}-\vec{G}')+V(\vec{G}-\vec{G}')-E(2\pi)^3\delta(\vec{G}-\vec{G}')\right)c(\vec{G}) = 0$$

$$V(\vec{G}) = \int d^3r\ V(\vec{r})\mathrm{e}^{i\vec{G}\vec{r}}$$

Thus we obtain the form

$$\left(\frac{\hbar^2\vec{G}'^2}{2m}-E\right)c(\vec{G}) + \int\frac{d^3G'}{(2\pi)^3}V(\vec{G}-\vec{G}')c(\vec{G}') = 0$$

We can describe this in the form that the potential transfers a momentum $\hbar(\vec{G}-\vec{G}')$ to an electron.

If the potential is periodic its Fourier transform is**Editor: Check**

$$V(\vec{G}) = \sum_{i,j,k}(2\pi)^3 V(\vec{G}_{i,j,k})\delta(\vec{G}-\vec{G}_{i,j,k})$$

Thus we obtain the form

$$\left(\frac{\hbar^2\vec{G}'^2}{2m}-E\right)c(\vec{G}) + \sum_{i,j,k}V(\vec{G}_{i,j,k})c(\vec{G}+\vec{G}_{i,j,k}) = 0$$

If we make the Ansatz $c(\vec{G}) = c_{\vec{k},i,j,k}\delta\left(\vec{G} - \left(\vec{k} + \sum_{i,j,k}\vec{G}_{i,j,k}\right)\right)$ we immediately observe that the Schrödinger equation can be solved independently for each set of coefficients with different $\vec{k}$.

Thus it makes sense to draw the energies of the basis states for each $\vec{k}$ on top of each other. Thus we obtain the **reduced zone scheme**. If we again fold out the states, we obtain the **extended zone scheme**.

We will see later in the chapter on first order perturbation theory, that the energies that lie close to each other "repel" each other. This repulsion is strongest if the energy spacing is small. This leads to an opening of the band gap wherever two bands in the reduced zone scheme meet.

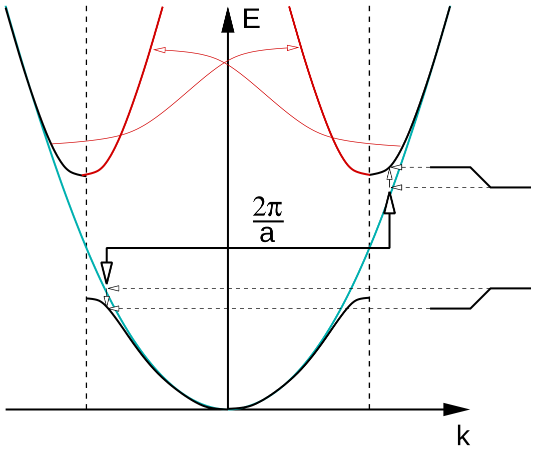


Fig. 10.5: If a free particle experiences a potential with periodicity a, states with wave vectors that differ by a reciprocal lattice vector $G_n = \frac{2\pi}{a}n$ couple. The coupling shifts the upper state up and the lower state down, which results in an opening of the gap. As a result of the coupling, the wave vector of a free particle state is no more a good quantum number. However, the wave vector in the reduced zone scheme remains to be a good quantum number.

10.6 Some common Lattices

10.6.1 Cubic lattices

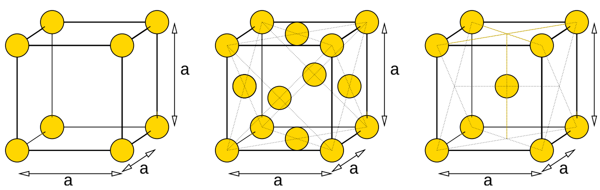


Fig. 10.6: The simple-cubic, the face-centered cubic and the body-centered cubic lattice in the simple cubic unit cell.

The most simple lattice is the simple cubic (sic) lattice which has lattice vectors

$$\vec{T}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} a \quad , \quad \vec{T}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} a \quad , \quad \vec{T}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} a$$

where a is the lattice constant. I am not aware of any crystal with one atom per unit cell that is stable in the simple cubic lattice.

More common is the face-centered cubic (fcc) lattice with lattice vectors

$$\vec{T}_1 = \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} a \quad , \qquad \vec{T}_2 = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} a \quad , \qquad \vec{T}_3 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} a$$

The face centered cubic lattice corresponds to one of the possible lattices of closed packed spheres. It is very common for atoms which do not form directed bonds. One example are noble gases, when they freeze. Another are metals such as Sodium, Aluminium, Gold, Copper, Silver. Austenite, the major component of Steel is another example for a fcc-crystal.

The third cubic lattice is the body-centered cubic (bcc) lattice with lattice vectors

$$\vec{T}_1 = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} a \quad , \qquad \vec{T}_2 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} a \quad , \qquad \vec{T}_3 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} a$$

The body centered cubic lattice is found for metals, that form some extend of directed bonds, such as the transition metals in the middle of the transition metal row. One example is iron in its magnetic form

10.7 Recommended exercises

1. Symmetrization of states with a discrete symmetry: Exercise 16.6 on p. 246 (No worked solutions)
2. Kronig-Penney model: Exercise 16.7 on p. 247 is an application of the Bloch theorem.

Chapter 11

Rotations, Angular Momentum and Spin

Systems with rotational symmetry play such an important role, that we devote a separate chapter to them, even though what is done here is only an extension of what we learned in the section about symmetry.

11.1 Rotations

Let us consider first a continuous rotation about the z-axis. The symmetry operator transforms $\vec{r}$ into $\vec{r'} = \boldsymbol{R}(\varphi)\vec{r}$, where $\boldsymbol{R}(\varphi)$ is the corresponding rotation matrix.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \underbrace{\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}}_{\boldsymbol{R}} \begin{pmatrix} x \\ y \end{pmatrix}$$

The angle is positive for counterclockwise rotation.

A rotation of a wave function by an angle φ is then defined as

$$\hat{S}(\varphi) \stackrel{\text{def}}{=} \int d^3r \; |\boldsymbol{R}(\varphi)\vec{r}\rangle\langle\vec{r}| \tag{11.1}$$

$$\underbrace{\langle\vec{r}|\hat{S}(\varphi)|\psi\rangle}_{\hat{S}(\varphi)\psi(\vec{r})} = \int d^3r \; \langle\vec{r}|\boldsymbol{R}\vec{r}\rangle\langle\vec{r}|\psi\rangle = \int d^3r \; \langle\boldsymbol{R}^{-1}\vec{r}|\vec{r}\rangle\langle\vec{r}|\psi\rangle = \underbrace{\langle\boldsymbol{R}^{-1}\vec{r}|\psi\rangle}_{\psi(\boldsymbol{R}^{-1}\vec{r})}$$

Note, that we need to use the inverse rotation, because we obtain the value of the rotated wave function at a point r_0 by looking up the value of the original wave function at the position which fell into r_0 by the rotation. The inverse of the rotation is simply its transpose, because the rotation matrix is unitary.

11.1.1 Derivation of the Angular momentum

Let us construct the canonical conjugate momentum L to the angle φ. We will see that L is the angular momentum.

$$
\begin{aligned}
\hat{L}_z\psi(x,y) &= i\hbar\lim_{\varphi\to 0}\Big(\frac{\hat{S}(\varphi)-\hat{1}}{\varphi}\Big)\psi(x,y)\\
&= i\hbar\lim_{\varphi\to 0}\Big(\frac{\psi[\overbrace{\cos(\varphi)}^{1+O(\varphi^2)}x+\overbrace{\sin(\varphi)}^{\varphi+O(\varphi^2)}y,-\overbrace{\sin(\varphi)}^{\varphi+O(\varphi^2)}x+\overbrace{\cos(\varphi)}^{1+O(\varphi^2)}y]-\psi[x,y]}{\varphi}\Big)\\
&= i\hbar\lim_{\varphi\to 0}\Big(\frac{\overbrace{\psi[x+\varphi y+O(\varphi^2),y-\varphi x+O(\varphi^2)]}^{\psi(x,y)+\big(x+\varphi y+O(\varphi)^2\big)\partial_x\psi+\big(y-\varphi x+O(\varphi)^2\big)\partial_y\psi+\dots}\;-\psi[x,y]}{\varphi}\Big)\\
&= i\hbar\lim_{\varphi\to 0}\Big(\frac{\varphi y\partial_x\psi-\varphi x\partial_y\psi+O(\varphi^2)}{\varphi}\Big)\\
&= \frac{\hbar}{i}\Big(-y\partial_x+x\partial_y\Big)\psi\\
\hat{L}_z &= \hat{x}\hat{p}_y-\hat{y}\hat{p}_x
\end{aligned}
$$

Thus we obtained the form of the expression of the angular-momentum operator.

Editor: Use the above example as exercise for a Taylor expansion

We can generalize this result into three dimensions for a rotation about an axis, which is described by a vector $\vec{\varphi}$. The direction of the vector describes the rotation axis and the length of the vector denotes the angle of rotation.

Definition 11.1 *ANGULAR-MOMENTUM OPERATOR*
The **angular-momentum operator** $\hat{\vec{L}}$ *has the form.*

$$\hat{\vec{L}}=\vec{r}\times\vec{p} \tag{11.2}$$

With the angular-momentum operator we can express the rotation operator $\hat{S}(\varphi)$ as

ROTATION OPERATOR

$$\hat{S}(\varphi)\stackrel{\text{Eq. 11.1}}{=}\int d^3r\;|\boldsymbol{R}\vec{r}\rangle\langle\vec{r}|\stackrel{\text{Eq. 9.7}}{=}\mathrm{e}^{-\frac{i}{\hbar}\hat{L}_x\varphi} \tag{11.3}$$

11.1.2 Another derivation of the angular momentum (Home study)

This derivation is more complex but may also show some aspects more clearly. It does not refer to wave functions and rests entirely within Dirac's bracket notation. In any case it is a very good exercise for the latter.

$$
\begin{aligned}
\hat{L}_z &= i\hbar\lim_{\varphi\to 0}\frac{1}{\varphi}\big[\hat{S}(\varphi)-\hat{1}\big]=i\hbar\lim_{\varphi\to 0}\frac{1}{\varphi}\Big[\int d^3r|\boldsymbol{R}(\varphi)\vec{r}\rangle\langle\vec{r}|-\int d^3r|\vec{r}\rangle\langle\vec{r}|\Big]\\
&= i\hbar\int d^3r\lim_{\varphi\to 0}\frac{1}{\varphi}\left[|\boldsymbol{R}(\varphi)\vec{r}\rangle-|\vec{r}\rangle\right]\langle\vec{r}
\end{aligned}
$$

Next we Taylor-expand the rotated vector $\boldsymbol{R}(\varphi)\vec{r}$ up to first order in φ.

$$\begin{aligned}\boldsymbol{R}(\varphi) &= \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+O(\varphi^2) & -\varphi+O(\varphi^2) & 0 \\ \varphi+O(\varphi^2) & 1+O(\varphi^2) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \mathbf{1} + \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\varphi + O(\varphi^2) \\ \Rightarrow \boldsymbol{R}\vec{r} &= \vec{r} + \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}\varphi + O(\varphi^2) = \vec{r} + \varphi\vec{e}_z \times \vec{r} + O(\varphi)^2\end{aligned}$$

Let us now use the translation operator $\hat{T}(\vec{\Delta}) \stackrel{\text{def}}{=} \int d^3r |\vec{r}+\vec{\Delta}\rangle\langle\vec{r}|$ to express the rotated state

$$\begin{aligned}|\boldsymbol{R}(\varphi)\vec{r}\rangle &= |\vec{r} + \underbrace{\varphi\vec{e}_z \times \vec{r} + O(\varphi^2)}_{\Delta}\rangle = \hat{T}(\underbrace{\varphi\vec{e}_z \times \vec{r} + O(\varphi^2)}_{\Delta})|\vec{r}\rangle \\ &\stackrel{\text{Eq. 9.7}}{=} \underbrace{\mathrm{e}^{-\frac{i}{\hbar}\hat{\vec{p}}\overbrace{\left(\varphi\vec{e}_z \times \vec{r} + O(\varphi^2)\right)}^{\Delta}}}_{\hat{T}(\Delta)}|\vec{r}\rangle \\ &\stackrel{Taylor}{=} \left[\hat{1} - \frac{i}{\hbar}\hat{\vec{p}}\varphi(\vec{e}_z \times \vec{r}) + O(\varphi^2)\right]|\vec{r}\rangle \\ &\stackrel{\vec{a}(\vec{b}\times\vec{c})=\vec{b}(\vec{c}\times\vec{a})}{=} \left[\hat{1} - \frac{i}{\hbar}\vec{e}_z(\vec{r} \times \hat{\vec{p}})\varphi + O(\varphi^2)\right]|\vec{r}\rangle\end{aligned}$$

Thus we obtain

$$\begin{aligned}\hat{L}_z &= i\hbar\int d^3r\, \underbrace{\lim_{\varphi\to 0}\left(-\frac{i}{\hbar}\vec{e}_z(\vec{r}\times\hat{\vec{p}})\varphi + O(\varphi^2)\right)}_{\frac{1}{i\hbar}\vec{e}_z(\vec{r}\times\hat{\vec{p}})}|\vec{r}\rangle\langle\vec{r}| = \vec{e}_z(\underbrace{\int d^3r\, |\vec{r}\rangle\vec{r}\langle\vec{r}|}_{\hat{\vec{r}}} \times \hat{\vec{p}}) \\ &= \vec{e}_z(\hat{\vec{r}} \times \hat{\vec{p}})\end{aligned}$$

Since we can choose our coordinate system arbitrarily we can immediately generalize the result directly to arbitrary directions of the rotation axis, which identifies the angular-momentum operator as

$$\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}}$$

11.2 Commutator Relations

It turns out that the components of the angular momentum operator do not commutate with each other. Therefore, it is not possible to determine the direction of the angular momentum with certainty. Let us now explore the commutator relations among the angular momentum operators.

$$\begin{aligned}[\hat{L}_x, \hat{L}_y] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] \\ &= y\hat{p}_z\hat{z}\hat{p}_x - \hat{y}\hat{p}_z\hat{x}\hat{p}_z - \hat{z}\hat{p}_y\hat{z}\hat{p}_x + \hat{z}\hat{p}_y\hat{x}\hat{p}_z \\ &\quad - \hat{z}\hat{p}_x\hat{y}\hat{p}_z + \hat{z}\hat{p}_x\hat{z}\hat{p}_y + \hat{x}\hat{p}_z\hat{y}\hat{p}_z - \hat{x}\hat{p}_z\hat{z}\hat{p}_y \\ &= \hat{y}\hat{p}_x[\hat{p}_z, \hat{z}] + \hat{p}_y\hat{x}[\hat{z}, \hat{p}_z] = i\hbar(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) \\ &= i\hbar\hat{L}_z\end{aligned}$$

The other commutators are obtained by cyclic permutation of indices.

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z; \qquad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x; \qquad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y$$

Furthermore we can show that $\hat{\vec{L}}^2$ commutates with all components

$$
\begin{aligned}
[\hat{\vec{L}}^2, \hat{L}_x] &= \underbrace{[\hat{L}_x^2, \hat{L}_x]}_{=0} + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x] \\
&= \Big(\hat{L}_y^2\hat{L}_x \overbrace{-\hat{L}_y\hat{L}_x\hat{L}_y\Big) + \Big(\hat{L}_y\hat{L}_x\hat{L}_y}^{=0} - \hat{L}_x\hat{L}_y^2\Big) + \Big(\hat{L}_z^2\hat{L}_x \overbrace{-\hat{L}_z\hat{L}_x\hat{L}_z\Big) + \Big(\hat{L}_z\hat{L}_x\hat{L}_z}^{=0} - \hat{L}_x\hat{L}_z^2\Big) \\
&= \hat{L}_y[\hat{L}_y, \hat{L}_x] + [\hat{L}_y, \hat{L}_x]\hat{L}_y + \hat{L}_z[\hat{L}_z, \hat{L}_x] + [\hat{L}_z, \hat{L}_x]\hat{L}_z \\
&= i\hbar\Big(-\hat{L}_y\hat{L}_z - \hat{L}_z\hat{L}_y + \hat{L}_z\hat{L}_y + \hat{L}_y\hat{L}_z\Big) = 0
\end{aligned}
$$

Cyclic permutation of indices yields the commutator relations with the remaining angular momentum components. Thus we can measure the amplitude of the angular momentum and one component of the angular momentum simultaneously.

	$[\hat{L}_i, \hat{L}_x]$	$[\hat{L}_i, \hat{L}_y]$	$[\hat{L}_i, \hat{L}_z]$	$[\hat{L}_i, \hat{\vec{L}}^2]$
$[\hat{L}_x, \hat{L}_i]$	0	$i\hbar\hat{L}_z$	$-i\hbar\hat{L}_y$	0
$[\hat{L}_y, \hat{L}_i]$	$-i\hbar\hat{L}_z$	0	$i\hbar\hat{L}_x$	0
$[\hat{L}_z, \hat{L}_i]$	$i\hbar\hat{L}_y$	$-i\hbar\hat{L}_x$	0	0

In short, we can write the result using the Levi-Civita Symbol[1] as

COMMUTATOR RELATION OF ANGULAR MOMENTA

$$[\hat{L}_i, \hat{L}_j] = i\hbar \sum_k \epsilon_{i,j,k}\hat{L}_k \tag{11.4}$$

11.3 Algebraic derivation of the eigenvalue spectrum

In the following we will show the algebraic description of angular momentum, which will provide us with the eigenvalues of angular momenta. The algebraic description is important, because we will find eigenstates, that cannot be represented by a single-valued wave function. Any Ansatz of a real space wave function will therefore fail to produce those eigenstates.

Shift operators revisited

We already made contact with shift operators in connection with continuous translations. There, it was possible use a shift operator to create new eigenstates from an existing eigenstate with a shifted eigenvalue. Here we use shift operators, that are not unitary operators and therefore are no symmetry transformations. However, we can still use them to create new eigenstates of an operator. The concept of shift operators to create new eigenstates is analogous to that of the creation and

[1] The Levi-Civita symbol is the fully antisymmetric tensor defined as being equal to one if the indices are ascending $\epsilon_{1,2,3} = \epsilon_{2,3,1} = \epsilon_{3,1,2} = 1$ am being -1 if the indices are descending, that is $\epsilon_{3,2,1} = \epsilon_{1,3,2} = \epsilon_{2,1,3} = -11$. If two indices are equal the Levi-Civita symbol vanishes. The Levi-Civita Symbol is used to describe the cross-product $(\vec{a}\times\vec{b})_i = \sum_{j,k}\epsilon_{i,j,k}a_jb_k$.

annihilation operators, that have been used in the algebraic treatment of the harmonic oscillator and the two state system.

A shift operator $\hat{A}$ has a commutator with the shifted operator $\hat{B}$, which is proportional to the shift operator itself. As a result, the shift operator produces a new eigenstate $\hat{A}|\psi_b\rangle$ of $\hat{B}$ from an existing one, $|\psi_b\rangle$ with a shifted eigenvalue, that is

$$[\hat{A},\hat{B}] = c\hat{A} \qquad \text{and} \qquad \hat{B}|\psi_b\rangle = |\psi_b\rangle b \qquad \Rightarrow \qquad \hat{B}\Big(\hat{A}|\psi_b\rangle\Big) = \Big(\hat{A}|\psi_b\rangle\Big)(b-c)$$

We can confirm this easily:

$$\hat{B}\hat{A}|\psi_b\rangle = \hat{A}\underbrace{\hat{B}|\psi_b\rangle}_{|\psi_b\rangle b} - c\hat{A}|\psi\rangle$$
$$\hat{B}\Big(\hat{A}|\psi_b\rangle\Big) = \Big(\hat{A}|\psi_b\rangle\Big)(b-c)$$

Furthermore we can show that $[\hat{A}^\dagger,\hat{B}] = -c\hat{A}^\dagger$ if $\hat{B}$ is hermitian and c is a real number.

$$[\hat{A}^\dagger,\hat{B}] = \hat{A}^\dagger\hat{B} - \hat{B}\hat{A}^\dagger = \Big(\hat{B}^\dagger\hat{A} - \hat{A}\hat{B}^\dagger\Big)^\dagger \stackrel{B=B^\dagger}{=} -[\hat{A},\hat{B}]^\dagger \stackrel{c=c^*}{=} -c\hat{A}^\dagger$$

Hence we can use $\hat{A}^\dagger$ to create an eigenstate of $\hat{B}$ with eigenvalue $b+c$

$$\hat{B}\Big(\hat{A}^\dagger|\Psi\rangle\Big) = \Big(\hat{A}^\dagger|\Psi\rangle\Big)(b+c)$$

Thus we can produce a whole series of eigenstates of $\hat{B}$ with equi-spaced eigenvalues, if we can only find an operator with the property $[\hat{A},\hat{B}] = c\hat{A}$, where c is real. The series of eigenvalues is bound from below and above if there is a solution to $\hat{A}|\psi_0\rangle = 0$ and to $\hat{A}^{\dagger(n+1)}|\psi_0\rangle = 0$.

Shift operators for $\hat{L}_z$

Such a shift operator $\hat{A}$ is a handy thing to have. Let us look for one that may be useful for angular momentum operators. We consider

$$[\hat{L}_y,\hat{L}_z] = i\hbar\hat{L}_x$$
$$[\hat{L}_z,\hat{L}_x] = i\hbar\hat{L}_y$$

which suggests that an operator $\hat{L}_- = a\hat{L}_x + b\hat{L}_y$ may do the job. If we use this as Ansatz for $[\hat{L}_-,\hat{L}_z] = c\hat{L}_-$, we find one solution[2] with $a=1$, $b=-i$ and $C=\hbar$.

Thus we define the **shift operators**

$$\hat{L}_- = \hat{L}_x - i\hat{L}_y \tag{11.5}$$
$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y = \hat{L}_-^\dagger \tag{11.6}$$

They obey the commutation relations

$$[\hat{L}_-,\hat{L}_z] = \hbar\hat{L}_-; \qquad [\hat{L}_+,\hat{L}_z] = -\hbar\hat{L}_+; \qquad [\hat{L}_-,\hat{\vec{L}}^2] = 0; \qquad [\hat{L}_+,\hat{\vec{L}}^2] = 0$$

[2]We use the ansatz $\hat{L}_- = a\hat{L}_x + b\hat{L}_y$ and require

$$[\hat{L}_-,\hat{L}_z] = C\hat{L}_-$$
$$\Rightarrow \qquad [a\hat{L}_x + b\hat{L}_y,\hat{L}_z]_- = a[\hat{L}_x,\hat{L}_z]_- + b[\hat{L}_y,\hat{L}_z]_- = -i\hbar a\hat{L}_y + i\hbar b\hat{L}_x = \left(ib\hbar\hat{L}_x - ia\hbar\hat{L}_y\right) \stackrel{!}{=} C\left(a\hat{L}_x + b\hat{L}_y\right)$$

we obtain two conditions, namely $aC = ib\hbar$ and $bC = -ia\hbar$. Thus $C = i\hbar\frac{b}{a} = -i\hbar\frac{a}{b}$, which requires $b^2 = -a^2$ and therefore $b = \pm ia$. If we choose $a=1$ and $b=-ia$, then $C=\hbar$. The other solution leads to L_+ with an arbitrary scale factor.

The first two equations identify the operators $\hat{L}_+$ and $\hat{L}_-$ as shift operators for eigenstates of $\hat{L}_z$. Thus for an eigenstate $|\mu\rangle$ of $\hat{L}_z$ with eigenvalue μ, the state $\hat{L}_+|\mu\rangle$ is an eigenstate of $\hat{L}_z$ with eigenvalue $\mu+\hbar$. Similarly the state $\hat{L}_-|\mu\rangle$ is an eigenstate of $\hat{L}_z$ with eigenvalue $\mu-\hbar$, i.e.

$$\hat{L}_z|\mu\rangle = |\mu\rangle\mu \qquad \Rightarrow \qquad \begin{cases} \hat{L}_z \underbrace{\hat{L}_+|\mu\rangle}_{|\mu+\hbar\rangle} = \underbrace{\hat{L}_+|\mu\rangle}_{|\mu+\hbar\rangle}(\mu+\hbar) \\ \hat{L}_z \underbrace{\hat{L}_-|\mu\rangle}_{|\mu-\hbar\rangle} = \underbrace{\hat{L}_-|\mu\rangle}_{|\mu-\hbar\rangle}(\mu-\hbar) \end{cases} \tag{11.7}$$

We can now express $\hat{\vec{L}}^2$ by $\hat{L}^+$, $\hat{L}_-$ and $\hat{L}_z$, which we will need later.

$$\hat{L}_+\hat{L}_- = (\hat{L}_x + i\hat{L}_y)(\hat{L}_x - i\hat{L}_y) = \underbrace{\hat{L}_x^2 + \hat{L}_y^2}_{\vec{L}^2 - L_z^2} - i\underbrace{[\hat{L}_x, \hat{L}_y]}_{i\hbar L_z}$$

$$= \hat{\vec{L}}^2 - \hat{L}_z^2 + \hbar\hat{L}_z \tag{11.8}$$

$$\hat{L}_-\hat{L}_+ = \hat{\vec{L}}^2 - \hat{L}_z^2 - \hbar\hat{L}_z \tag{11.9}$$

$$\hat{\vec{L}}^2 = \hat{L}_z^2 + \frac{1}{2}(\hat{L}_-\hat{L}_+ + \hat{L}_+\hat{L}_-) \tag{11.10}$$

Angular-momentum eigenvalue spectrum

Now we are ready to explore the spectrum of eigenvalues. Let us now assume we have eigenstates[3] $|\lambda,\mu\rangle$ to $\hat{\vec{L}}^2$ and $\hat{L}_z$, so that

$$\hat{\vec{L}}^2|\lambda,\mu\rangle = |\lambda,\mu\rangle\lambda \tag{11.11}$$

$$\hat{L}_z|\lambda,\mu\rangle = |\lambda,\mu\rangle\mu. \tag{11.12}$$

$$\langle\lambda,\mu|\lambda',\mu'\rangle = \delta_{\lambda,\lambda'}\delta_{\mu,\mu'} \tag{11.13}$$

The requirement that the norm of $\hat{L}_\pm|\lambda,\mu\rangle$ is positive provides us with bounds for the spectrum of eigenvalues of $\hat{L}_z$. This, somewhat involved, derivation will be shown in the following:

Let us consider an eigenstate $|\lambda,\mu\rangle$ of $\hat{\vec{L}}^2$ and $\hat{L}_z$. We form two new states, namely

$$|\phi^+_{\lambda,\mu}\rangle \stackrel{\text{def}}{=} \hat{L}_+|\lambda,\mu\rangle$$
$$|\phi^-_{\lambda,\mu}\rangle \stackrel{\text{def}}{=} \hat{L}_-|\lambda,\mu\rangle$$

by operating once with $\hat{L}_+$ and once with $\hat{L}_-$ on it. Because the scalar product of any state must be positive, we obtain two inequalities.

$$\begin{aligned} 0 &\le \langle\phi^+_{\lambda,\mu}|\phi^+_{\lambda,\mu}\rangle \\ &= \langle\lambda,\mu|\hat{L}_+^\dagger\hat{L}_+|\lambda,\mu\rangle \\ &\stackrel{\hat{L}_+^\dagger=\hat{L}_-}{=} \langle\lambda,\mu|\hat{L}_-\hat{L}_+|\lambda,\mu\rangle \\ &\stackrel{\text{Eq. }11.9}{=} \langle\lambda,\mu|\hat{\vec{L}}^2 - \hat{L}_z^2 - \hbar\hat{L}_z|\lambda,\mu\rangle \\ &\stackrel{\text{Eq. }11.11,11.12,11.13}{=} \lambda - \mu^2 - \hbar\mu \\ \Rightarrow \mu^2 + \hbar\mu &\le \lambda \end{aligned} \tag{11.14}$$

[3] Note that $|\lambda,\mu\rangle$ should be read as $|\Psi_{\lambda,\mu}\rangle$. The symbol Ψ seems superfluous– it is the same symbol for all members in the specified set of states– it is normally dropped. The new notation just includes the indices in the ket symbol. For example we would write $\Psi_{\lambda\mu}(\vec{r}) = \langle\vec{r}|\lambda,\mu\rangle$

The equal sign holds exactly when $|\phi^+_{\lambda,\mu}\rangle$ is the zero state.

Completely analogously we obtain the second inequality from the condition that the scalar product of $|\phi^-_{\lambda,\mu}\rangle$ with itself is positive.

$$
\begin{aligned}
0 &\leq \langle\phi^-_{\lambda,\mu}|\phi^-_{\lambda,\mu}\rangle \\
&= \langle\lambda,\mu|\hat{L}^\dagger_-\hat{L}_-|\lambda,\mu\rangle \\
&\stackrel{\hat{L}^\dagger_-=\hat{L}_+}{=} \langle\lambda,\mu|\hat{L}_+\hat{L}_-|\lambda,\mu\rangle \\
&\stackrel{\text{Eq. 11.8}}{=} \langle\lambda,\mu|\hat{\vec{L}}^2-\hat{L}_z^2+\hbar\hat{L}_z|\lambda,\mu\rangle \\
&\stackrel{Eqs.\ 11.11,11.12,11.13}{=} \lambda-\mu^2+\hbar\mu \\
\Rightarrow \mu^2-\hbar\mu &\leq \lambda
\end{aligned}
\qquad (11.15)
$$

Again, the equal sign holds exactly when $|\phi^-_{\lambda,\mu}\rangle$ is the zero state.

The two inequalities are graphically shown in figure 11.1. Let us assume that there is a state

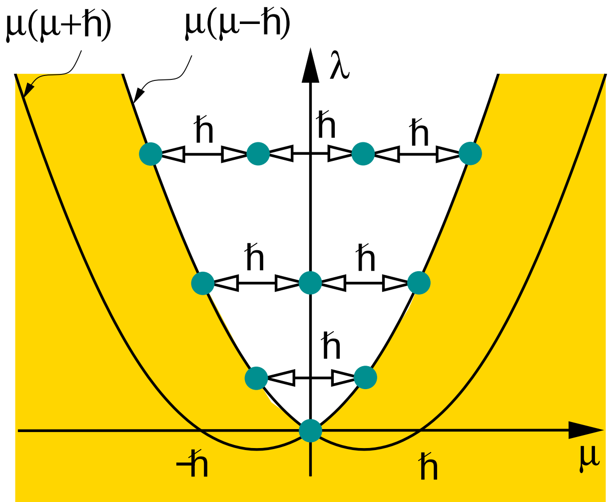


Fig. 11.1: Demonstration of the allowed values for λ and μ. The inequalities Eq. 11.14 and Eq. 11.15 require that the allowed values of (λ,μ) lie in the white region above both parabolas. The ladder of μ values terminates towards the right only on the left parabola, for which $\hat{L}_+|\lambda,\mu\rangle = |\varnothing\rangle$. Similarly it terminates towards the left only on the right parabola. For a bound set of μ values each ladder must start and end at the boundary of the white region. The discrete spacing, Eq. 11.7, of μ values by $\hbar$ specifies the allowed values of μ.

$|\lambda,\mu\rangle$ in the allowed region, that is both inequalities are fulfilled. This state is represented by a point in the white region of the diagram above. We have shown that each application of $\hat{L}_-$ creates a new state with μ reduced by $\hbar$. This creates a new point shifted to the left by $\hbar$. If we repeatedly apply $\hat{L}_-$ we create a point in the forbidden region. This however would contradict the inequalities. However, if we exactly hit the boundary, we are fine, because the next state is a zero state. That is, there is no next state, and therefore there is no contradiction. This consideration tells us that the only allowed states $|\lambda,\mu\rangle$ must be located a multiple of $\hbar$ to the right of the left boundary.

Now let us perform the same steps but we operate with $\hat{L}_+$ on our state $|\lambda,\mu\rangle$. $\hat{L}_+$ shifts the point in the graph by $\hbar$ to the right. The contradiction given by the inequalities can only be avoided if the horizontal distance between left and right boundary is exactly a multiple of $\hbar$. This condition immediately provides us with the allowed values for λ.

We will now express the same by formulas in order to determine the allowed values for λ and μ:

The inequalities Eq. 11.14 and Eq. 11.15 require $\mu_{min} \le \mu \le \mu_{max}$ with

$$\begin{aligned}
\mu^2_{min/max} \pm \hbar\,\mu_{min/max} - \lambda &\overset{\text{Eqs. }11.14,11.15}{=} 0 \\
\Rightarrow \left(\mu_{min/max} \pm \frac{\hbar}{2}\right)^2 &= \left(\frac{\hbar}{2}\right)^2 + \lambda \\
\Rightarrow \mu_{min/max}(\lambda) &= \pm\left(\frac{\hbar}{2} \mp \sqrt{\left(\frac{\hbar}{2}\right)^2 + \lambda}\right)
\end{aligned}$$

The choice of the second sign $\mp$ determines, whether we choose the inner or the outer branches of the two parabola in Fig. 11.1. Because λ is positive, as seen in Fig. 11.1, the negative sign yields the smaller absolute value for $\mu_{min/max}$.

Thus, the the lowest allowed value μ_{min} of μ and the highest allowed value μ_{max} of μ for a given λ is

$$\begin{aligned}
\mu_{min}(\lambda) &= +\left(\frac{\hbar}{2} - \sqrt{(\frac{\hbar}{2})^2 + \lambda}\right) \\
\mu_{max}(\lambda) &= -\left(\frac{\hbar}{2} - \sqrt{(\frac{\hbar}{2})^2 + \lambda}\right)
\end{aligned} \tag{11.16}$$

Note, that the value of the expression in parenthesis is negative.

The requirement that these values differ by a multiple of $\hbar$, obtained from Eq. 11.7, yields

$$\mu_{min} - \mu_{max} = \hbar n$$

where n is an arbitrary non-negative integer. Thus we obtain

$$\begin{aligned}
\underbrace{-\frac{\hbar}{2} + \sqrt{(\frac{\hbar}{2})^2 + \lambda}}_{\mu_{max}(\lambda)}\underbrace{-\frac{\hbar}{2} + \sqrt{(\frac{\hbar}{2})^2 + \lambda}}_{-\mu_{min}(\lambda)} &= -\hbar + \sqrt{\hbar^2 + 4\lambda} = \hbar n \\
\hbar^2(n+1)^2 &= \hbar^2 + 4\lambda \\
\lambda &= \hbar^2\frac{(n+1)^2 - 1}{4} = \hbar^2\frac{n^2 + 2n}{4} = \hbar^2[(\frac{n}{2})^2 + \frac{n}{2}] \\
\lambda &= \hbar^2\ell(\ell+1)
\end{aligned}$$

where we replaced $n/2$ by ℓ. ℓ can assume half-integer values $\ell = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$. Thus we have obtained the eigenvalue spectrum of $\hat{\vec{L}}^2$.

The eigenvalues for $\hat{L}_z$ are obtained for each ℓ from

$$\begin{aligned}
\mu_{min} &= \frac{\hbar}{2} - \sqrt{(\frac{\hbar}{2})^2 + \lambda} = \frac{\hbar}{2} - \sqrt{(\frac{\hbar}{2})^2 + \hbar^2\ell(\ell+1)} \\
&= \frac{\hbar}{2} - \sqrt{\hbar^2(\ell^2 + \ell + \frac{1}{4})} = \frac{\hbar}{2} - \sqrt{\hbar^2(\ell + \frac{1}{2})^2} = \frac{\hbar}{2} - \hbar(\ell + \frac{1}{2}) \\
&= -\hbar\ell
\end{aligned}$$

Now we change our notation. Instead of characterizing a state by its eigenvalues $\lambda = \hbar^2\ell(\ell+1)$ and $\mu = \hbar m$ as $|\lambda, \mu\rangle$, we use the angular quantum numbers ℓ and m as $|\ell, m\rangle$. Thus we denote the eigenstates of the angular momentum as $|\lambda, \mu\rangle = |\ell, m\rangle$.

EIGENVALUE SPECTRUM OF THE ANGULAR MOMENTUM

$$\hat{\vec{L}}^2|\ell, m\rangle = |\ell, m\rangle\hbar^2\ell(\ell+1)$$
$$\hat{L}_z|\ell, m\rangle = |\ell, m\rangle\hbar m$$

where the quantum numbers $\ell = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ can have any non-negative, half-integer value and $m = -\ell, -\ell+1, \ldots, \ell-1, \ell$.

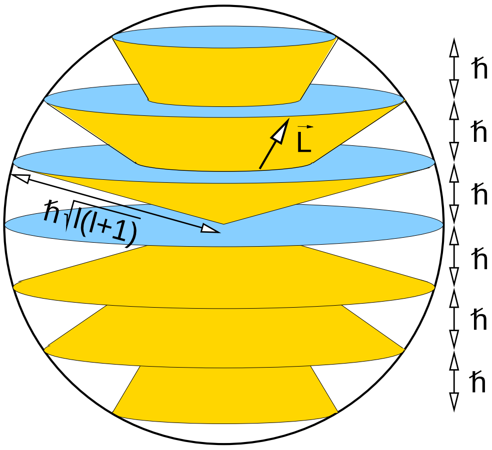


Fig. 11.2: The absolute value of the angular momentum can have the values $\hbar\sqrt{\ell(\ell+1)}$. The eigenvalues of $\hat{L}_z$ can assume the values $\hbar m$, where $m = -\ell, -\ell+1, \ldots, \ell-1, \ell$. In an eigenstate the values of $\hat{L}_z$ and $\hat{L}_y$ cannot be determined. It is as if the angular momentum would be precessing. However, the angular momentum does not move, but it simply is delocalized on the circumference of the circles shown. Note that $\hat{L}_x^2 + \hat{L}_y^2 = \hat{\vec{L}}^2 - \hat{L}_z^2$ does have sharp values.

Recursion relation for eigenstates

So far we have obtained the eigenvalues and a procedure to create new eigenstates for different eigenvalues of $\hat{L}_z$ if one eigenstate is known. However these states are not automatically normalized. Here we want to renormalize the shift operators so that they create normalized states from normalized states

$$\begin{aligned}
\langle \ell, m|\hat{L}_+^\dagger \hat{L}_+|\ell, m\rangle &= \langle \ell, m|\hat{\vec{L}}^2 - \hat{L}_z^2 - \hbar\hat{L}_z|\ell, m\rangle \\
&= \langle \ell, m|\ell, m\rangle(\hbar^2\ell(\ell+1) - \hbar^2 m^2 - \hbar^2 m) \\
&= \langle \ell, m|\ell, m\rangle\hbar^2(\ell(\ell+1) - m(m+1)) \\
|\ell, m+1\rangle &= \hat{L}_+|\ell, m\rangle\frac{1}{\hbar\sqrt{\ell(\ell+1) - m(m+1)}} \\
&= \hat{L}_+|\ell, m\rangle\frac{1}{\hbar\sqrt{(\ell-m)(\ell+m+1)}}
\end{aligned}$$

Similarly we obtain the renormalization for $\hat{L}_-$

$$\begin{aligned}
\langle \ell, m|\hat{L}_-^\dagger \hat{L}_-|\ell, m\rangle &= \langle \ell, m|\hat{\vec{L}}^2 - \hat{L}_z^2 + \hbar\hat{L}_z|\ell, m\rangle \\
&= \langle \ell, m|\ell, m\rangle(\hbar^2\ell(\ell+1) - \hbar^2 m^2 + \hbar^2 m) \\
&= \langle \ell, m|\ell, m\rangle\hbar^2(\ell(\ell+1) - m(m-1)) \\
|\ell, m-1\rangle &= \hat{L}_-|\ell, m\rangle\frac{1}{\hbar\sqrt{\ell(\ell+1) - m(m-1)}} \\
&= \hat{L}_-|\ell, m\rangle\frac{1}{\hbar\sqrt{(\ell+m)(\ell-m+1)}}
\end{aligned}$$

Thus if we know just one eigenstate of the angular momentum for each ℓ we can construct all others by the recursion relation above.

Let us summarize here the effect of the angular momentum operators on the basis of eigenstates of $\hat{\vec{L}}^2$ and $\hat{L}_z$:

$$\begin{aligned}
\hat{\vec{L}}^2|\ell, m\rangle &= |\ell, m\rangle\hbar^2\ell(\ell+1) && (11.17) \\
\hat{L}_z|\ell, m\rangle &= |\ell, m\rangle\hbar m && (11.18) \\
\hat{L}_+|\ell, m\rangle &= |\ell, m+1\rangle\hbar\sqrt{(\ell-m)(\ell+m+1)} && (11.19) \\
\hat{L}_-|\ell, m\rangle &= |\ell, m-1\rangle\hbar\sqrt{(\ell+m)(\ell-m+1)} && (11.20)
\end{aligned}$$

Because we can express all angular momentum components by $\hat{L}_z, \hat{L}_+, \hat{L}_-$, we can derive the matrix elements of any operator composed from angular momenta just from these four relations.[4]

11.4 Eigenstates of angular momentum: spherical harmonics

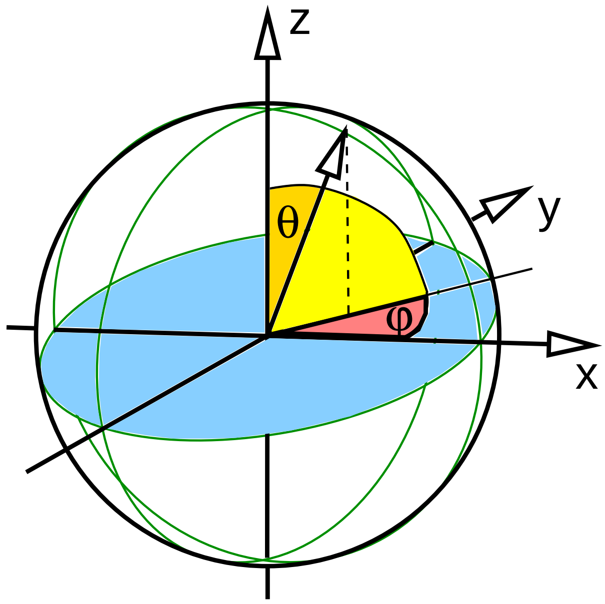
In the following we will derive the functional form of the eigenstates of $\hat{L}_z$ and $\hat{\vec{L}}^2$, which are called **spherical harmonics**.

Let us introduce polar coordinates:

$$\begin{aligned}
x &= r\cos(\phi)\sin(\theta) \\
y &= r\sin(\phi)\sin(\theta) \\
z &= r\cos(\theta)
\end{aligned}
\qquad\qquad
\begin{aligned}
r &= \sqrt{x^2+y^2+z^2} \\
\theta &= \mathrm{acos}(\frac{z}{r}) \\
\phi &= \mathrm{acos}(\frac{x}{\sqrt{x^2+y^2}})\mathrm{sgn}(y)
\end{aligned}$$

with $\theta \in [0, 2\pi]$ and $\phi \in [-\pi, \pi]$

[4] The first of these relations is actually redundant, and follows from the remaining three.



Using Eq. 9.7, we can express the rotation operator $\hat{S}(\alpha)$

$$\langle r,\theta,\phi|\hat{S}(\alpha)|\psi\rangle = \underbrace{\langle r,\theta,\phi-\alpha|\psi\rangle}_{\psi_m(r,\theta,\phi-\alpha)} \tag{11.21}$$

which shifts the angle by ϕ about the z-axis in the form

$$\hat{S}(\phi) \stackrel{\text{Eq. 11.3}}{=} \mathrm{e}^{-\frac{i}{\hbar}\hat{L}_z\phi} \tag{11.22}$$

If the shift operator is applied to an eigenstate $|\psi_m\rangle$ of the z-component of the angular momentum $\hat{L}_z$ with eigenvalue $\hbar m$, we obtain

$$\hat{S}(\alpha)|\psi_m\rangle \stackrel{\text{Eq. 11.22}}{=} \mathrm{e}^{-\frac{i}{\hbar}\hat{L}_z\alpha}|\psi_m\rangle \stackrel{\hat{L}_z|\psi_m\rangle=|\psi_m\rangle\hbar m}{=} |\psi_m\rangle\mathrm{e}^{-im\alpha}$$
$$\langle r,\theta,\phi|\hat{S}(\alpha)|\psi_m\rangle \stackrel{\text{Eq. 11.21}}{=} \langle r,\theta,\phi-\alpha|\psi_m\rangle = \langle r,\theta,\phi|\psi_m\rangle\mathrm{e}^{-im\alpha}$$

setting first $\phi=0$ and then $\alpha=-\phi$ we obtain

$$\psi_m(r,\theta,\phi) = \psi_m(r,\theta,0)\mathrm{e}^{im\phi}$$

Thus we obtained the ϕ-dependence of the angular momentum. This works fine for integer angular momentum. For a half-integer m, the wave function changes its sign for a full turn, because $\mathrm{e}^{i\pi}=-1$. Hence they cannot be associated with a single valued wave function if m is not an integer. Half-integer angular momenta are associated to spin, which we will discuss later.

The θ-dependence can be obtained from $\hat{L}_-Y_{\ell,-\ell}=0$ and $\hat{L}_+Y_{\ell,m}=Y_{\ell,m+1}\hbar\sqrt{\ell(\ell+1)-m(m-1)}$. Here we only summarize the main results.

The full eigenstates are[34]

$$Y_{\ell,m}(\theta,\phi) = \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}P_\ell^m(\cos\theta)\mathrm{e}^{im\phi} \tag{11.23}$$

The P_ℓ^m are called associated Legendre polynomials defined as

$$P_\ell^{|m|}(u) = (-1)^m\sqrt{1-u^2}^{|m|}\frac{d^{|m|}}{du^{|m|}}P_\ell(u). \tag{11.24}$$

$$P_\ell^{-|m|}(u) = (-1)^m\frac{(\ell-|m|)!}{(\ell+|m|)!}P_\ell^{|m|}(u) \tag{11.25}$$

with the help of the Legendre polynomials

$$P_\ell(u) = \frac{1}{2^\ell \ell!}\partial_u^\ell (u^2-1)^\ell \tag{11.26}$$

Numerically, associated Legendre polynomials are be evaluated by efficient recursion relations. Most recursion relations amplify numerical errors, rendering them useless. A recursion relation for associated Legendre polynomials that works is found in the "Numerical Recipes"[34]. It is in general advisable to consult a book on numerical mathematics such as the "Numerical Recipes", before evaluating special functions on the computer, as they often pose substantial difficulties that are easily overlooked.

	$Y_{\ell,m}(\vec{r})$	$Y_{\ell,m}(\theta,\phi)$
$Y_{0,0}$	$\sqrt{\frac{1}{4\pi}}$	$\sqrt{\frac{1}{4\pi}}$
$Y_{1,-1}$	$\sqrt{\frac{3}{8\pi}}\frac{x-iy}{r}$	$\sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\phi}$
$Y_{1,0}$	$\sqrt{\frac{3}{4\pi}}\frac{z}{r}$	$\sqrt{\frac{3}{4\pi}}\cos(\theta)$
$Y_{1,1}$	$-\sqrt{\frac{3}{8\pi}}\frac{x+iy}{r}$	$-\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\phi}$
$Y_{2,-2}$	$\sqrt{\frac{15}{32\pi}}\frac{(x^2-y^2)-ixy}{r^2}$	$\sqrt{\frac{15}{32\pi}}\sin^2\theta e^{-2i\phi}$
$Y_{2,-1}$	$\sqrt{\frac{15}{8\pi}}\frac{xz-iyz}{r^2}$	$\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{-i\phi}$
$Y_{2,0}$	$\sqrt{\frac{5}{16\pi}}\frac{3z^2-r^2}{r^2}$	$\sqrt{\frac{5}{16\pi}}(3\cos^2\theta-1)$
$Y_{2,1}$	$-\sqrt{\frac{15}{8\pi}}\frac{xz+iyz}{r^2}$	$-\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{i\phi}$
$Y_{2,2}$	$\sqrt{\frac{15}{32\pi}}\frac{(x^2-y^2)+ixy}{r^2}$	$\sqrt{\frac{15}{32\pi}}\sin^2\theta e^{2i\phi}$

Table 11.1: Spherical harmonics $Y_{\ell,m}$ in Cartesian and polar coordinates. Polar coordinates are defined as $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$ and $z = r\cos\theta$. Source: Jackson: Electrodynamics and Atkins[4].

Often the **real spherical harmonics** are used, which are fully real, but are not eigenstates of $\hat{L}_z$. We denote them with the symbol $\bar{Y}$. For $m = 0$ real and complex spherical harmonics are identical. For $m \neq 0$ we use the following definition.

$$\bar{Y}_{\ell,|m|} = \frac{1}{\sqrt{2}}\Big(Y_{\ell,|m|} + (-1)^m Y_{\ell,-|m|}\Big) = \sqrt{2}\mathrm{Re}[Y_{\ell,m}]$$
$$\bar{Y}_{\ell,-|m|} = -\frac{i}{\sqrt{2}}\Big(Y_{\ell,|m|} - (-1)^m Y_{\ell,-|m|}\Big) = \sqrt{2}\mathrm{Im}[Y_{\ell,m}]$$

Note that the real spherical harmonics are still eigenstates to $\hat{\vec{L}}^2$, but they are no eigenstates of $\hat{L}_z$

The following is a list of real spherical harmonics from which the complex spherical harmonics are readily obtained. Here the real spherical harmonics are identified by a running index $n = \ell^2 + \ell + m$

s	$\bar{Y}_1$	Y_0	$\sqrt{\frac{1}{4\pi}}$
p_x	$\bar{Y}_2$	$\frac{1}{\sqrt{2}}(Y_{1,1} - Y_{1,-1})$	$\sqrt{\frac{3}{4\pi}}\frac{x}{r}$
p_z	$\bar{Y}_3$	$Y_{1,0}$	$\sqrt{\frac{3}{4\pi}}\frac{z}{r}$
p_y	$\bar{Y}_4$	$\frac{-i}{\sqrt{2}}(Y_{1,1} + Y_{1,-1})$	$\sqrt{\frac{3}{4\pi}}\frac{y}{r}$
$d_{x^2-y^2}$	$\bar{Y}_5$	$\frac{1}{\sqrt{2}}(Y_{2,2} + Y_{2,-2})$	$\sqrt{\frac{15}{16\pi}}\frac{x^2-y^2}{r^2}$
d_{xz}	$\bar{Y}_6$	$\frac{1}{\sqrt{2}}(Y_{2,1} - Y_{2,-1})$	$\sqrt{\frac{60}{16\pi}}\frac{xz}{r^2}$
$d_{3z^2-r^2}$	$\bar{Y}_7$	$Y_{2,0}$	$\sqrt{\frac{5}{16\pi}}\frac{3z^2-r^2}{r^2}$
d_{yz}	$\bar{Y}_8$	$\frac{-i}{\sqrt{2}}(Y_{2,1} + Y_{2,-1})$	$\sqrt{\frac{60}{16\pi}}\frac{yz}{r^2}$
d_{xy}	$\bar{Y}_9$	$\frac{-i}{\sqrt{2}}(Y_{2,2} - Y_{2,-2})$	$\sqrt{\frac{60}{16\pi}}\frac{xy}{r^2}$

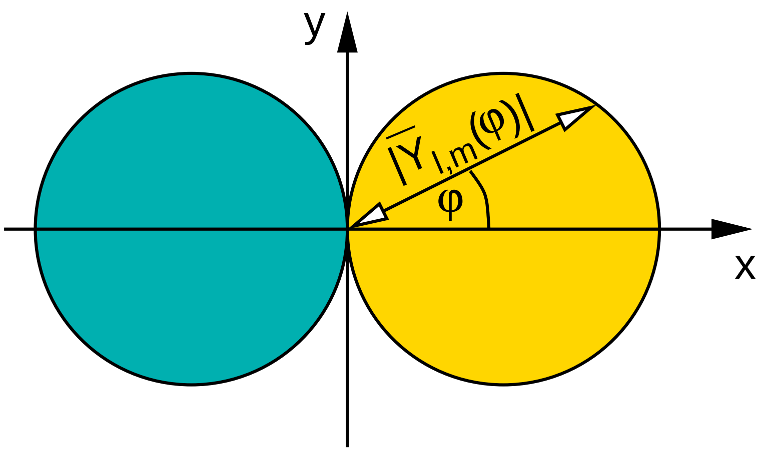


Fig. 11.3: Representation of real spherical harmonics ($\cos(\varphi)$) in polar coordinates. For a given direction defined by the angles θ, ϕ, a vector with a length equal to the absolute value of the amplitude of the real spherical harmonics points from the origin to the graph of the curve. Because this graph does not distinguish the different signs of the function, they are indicated by coloring the lobes with positive and negative values differently. Note that spherical harmonics defined as eigenstates of the angular momentum $\hat{\vec{L}}^2$ and $\hat{L}_z$ are complex, so that real and imaginary part of a spherical harmonics must be drawn separately. Real spherical harmonics are not eigenstates of $\hat{L}_z$, but they are already real.

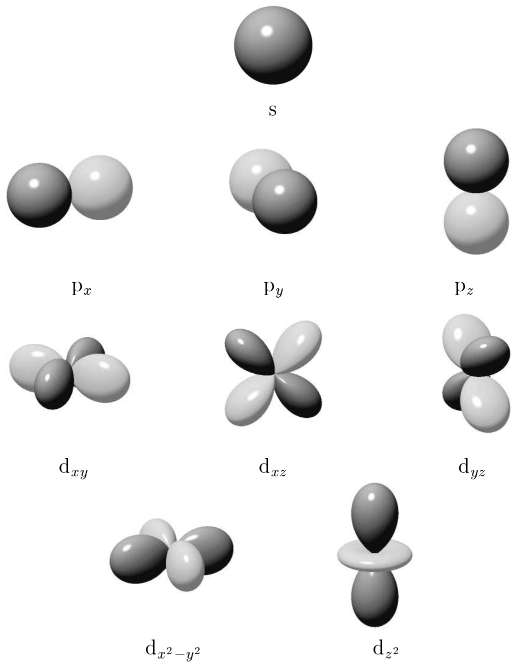


Fig. 11.4: Real spherical harmonics: The graphical representation refers to a surface $|\frac{1}{|r|}\bar{Y}_{\ell,m}(\vec{r})| = 1$. That is for every angle the distance of the surface from the origin is equal to the value of the spherical harmonic. Typically the lobes are colored differently to distinguish regions with positive and negative sign. The motivation for these plots is the representation of atomic orbitals, which are eigenstates of angular momentum and have the form $R(|r|)Y_{\ell,m}(\vec{r})$. If we plot a surface where the orbital has one value c in one color, and another surface for a value $-c$ with another color, we obtain similarly looking lobes.

11.5 Spin

An important result of the algebraic derivation of the spectrum of the angular momentum is the existence of states with half-integer m quantum number. There is no single valued wave function in real space that is an eigenstate to a non-integer m so that $\hat{L}_z|\psi\rangle = |\psi\rangle\hbar m$. Therefore, we could not have found these eigenvalues by trying to solve the differential equation in real space. The way out is to treat spin algebraically.

In the following we concentrate on eigenstates with $\ell = \frac{1}{2}$, which we also call spin $\frac{1}{2}$ eigenstates. States with other spin values ℓ can be represented analogously.

Let us introduce a basis of spin eigenstates, namely $|\uparrow\rangle \stackrel{\text{def}}{=} |\underbrace{\frac{1}{2}}_{\ell}, \underbrace{\frac{1}{2}}_{m}\rangle$ and $|\downarrow\rangle \stackrel{\text{def}}{=} |\frac{1}{2}, -\frac{1}{2}\rangle$ defined by

$$\hat{L}_z|\uparrow\rangle = |\uparrow\rangle\frac{\hbar}{2} \qquad \hat{L}_z|\downarrow\rangle = |\downarrow\rangle(-\frac{\hbar}{2})$$

so that

$$\begin{pmatrix} \langle\uparrow|\hat{L}_z|\uparrow\rangle & \langle\uparrow|\hat{L}_z|\downarrow\rangle \\ \langle\downarrow|\hat{L}_z|\uparrow\rangle & \langle\downarrow|\hat{L}_z|\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Now we use

$$\hat{L}_+|\ell, m\rangle = |\ell, m+1\rangle\hbar\sqrt{\ell(\ell+1) - m(m+1)}$$
$$\hat{L}_-|\ell, m\rangle = |\ell, m-1\rangle\hbar\sqrt{\ell(\ell+1) - m(m-1)}$$

with $\ell = \frac{1}{2}$ and $m = \pm\frac{1}{2}$ to express the action of the ladder operators $\hat{L}_+$ and $\hat{L}_-$ in terms of these basis functions

$$\hat{L}_-|\uparrow\rangle = |\downarrow\rangle\hbar; \qquad \hat{L}_-|\downarrow\rangle = 0$$
$$\hat{L}_+|\uparrow\rangle = 0; \qquad \hat{L}_+|\downarrow\rangle = |\uparrow\rangle\hbar$$

$$\begin{pmatrix} \langle\uparrow|\hat{L}_-|\uparrow\rangle & \langle\uparrow|\hat{L}_-|\downarrow\rangle \\ \langle\downarrow|\hat{L}_-|\uparrow\rangle & \langle\downarrow|\hat{L}_-|\downarrow\rangle \end{pmatrix} = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$
$$\begin{pmatrix} \langle\uparrow|\hat{L}_+|\uparrow\rangle & \langle\uparrow|\hat{L}_+|\downarrow\rangle \\ \langle\downarrow|\hat{L}_+|\uparrow\rangle & \langle\downarrow|\hat{L}_+|\downarrow\rangle \end{pmatrix} = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Using $\hat{L}_x = \frac{1}{2}(\hat{L}_+ + \hat{L}_-)$ and $\hat{L}_y = \frac{1}{2i}(\hat{L}_+ - \hat{L}_-)$ we obtain the remaining two components of the angular momentum

$$\begin{pmatrix} \langle\uparrow|\hat{L}_x|\uparrow\rangle & \langle\uparrow|\hat{L}_x|\downarrow\rangle \\ \langle\downarrow|\hat{L}_x|\uparrow\rangle & \langle\downarrow|\hat{L}_x|\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$\begin{pmatrix} \langle\uparrow|\hat{L}_y|\uparrow\rangle & \langle\uparrow|\hat{L}_y|\downarrow\rangle \\ \langle\downarrow|\hat{L}_y|\uparrow\rangle & \langle\downarrow|\hat{L}_y|\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Thus we can write

$$\hat{L}_i\begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix} = \frac{\hbar}{2}\sigma_i\begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix}$$

where $\Psi_\uparrow = \langle\uparrow|\Psi\rangle$ and $\Psi_\downarrow = \langle\downarrow|\Psi\rangle$ are vector components of a state $|\Psi\rangle$ in the basis of eigenstates to $\hat{L}_z$, and where the σ_i are the so-called **Pauli matrices**

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

that we have already learned about in the section on the two state system.

Most particles have an angular momentum that is not related to its “trajectory”. One interprets it as the rotation of a particle about its axis. A spin $\frac{1}{2}$ particle of this type is the electron. It is described by a double valued spatial wave function $(\Psi_\uparrow(r), \Psi_\downarrow(r))$, that is able to describe both its position and angular momentum coordinates.

The spin at a certain point is obtained from a two-component wave function by

$$S_i(r) = \frac{\hbar}{2}\begin{pmatrix} \Psi^*_\uparrow(r) \\ \Psi^*_\downarrow(r) \end{pmatrix}\sigma_i\begin{pmatrix} \Psi_\uparrow(r) \\ \Psi_\downarrow(r) \end{pmatrix}$$
$$S_x(r) = \frac{\hbar}{2}\Big(\Psi^*_\uparrow(r)\Psi_\downarrow(r) + \Psi^*_\downarrow(r)\Psi_\uparrow(r)\Big) = \hbar\mathrm{Re}[\Psi^*_\uparrow(r)\Psi_\downarrow(r)]$$
$$S_y(r) = \frac{\hbar}{2i}\Big(\Psi^*_\uparrow(r)\Psi_\downarrow(r) - \Psi^*_\downarrow(r)\Psi_\uparrow(r)\Big) = \hbar\mathrm{Im}[\Psi^*_\uparrow(r)\Psi_\downarrow(r)]$$
$$S_z(r) = \frac{\hbar}{2}\Big(\Psi^*_\uparrow(r)\Psi_\uparrow(r) - \Psi^*_\downarrow(r)\Psi_\downarrow(r)\Big)$$

The particle density is obtained as the sum of the squared components.

$$n(r) = \left(\Psi^*_\uparrow(r)\Psi_\uparrow + \Psi^*_\downarrow(r)\Psi_\downarrow(r)\right)$$

Often the wave functions are constrained to have exclusively either only a spin-up component or exclusively a spin-down component. Then the states are eigenstates to $S_z = \frac{\hbar}{2}\sigma_z$.

11.6 Addition of angular momenta

Let us consider two electrons orbiting in an atom. Obviously this is a rotationally invariant problem, and therefore the angular momentum must be a quantum number. However, if the particles interact, they can exchange angular momentum between themselves, as only the total angular momentum is conserved.

We can describe the two electrons by their individual angular momenta $\vec{L}_1$ and $\vec{L}_2$, where $\vec{L}_1$ acts only on the coordinates and momenta of the first electron and $\vec{L}_2$ acts only on the coordinates and momenta of the second atom. **Examples:**

- For example if the two electrons are described by a two-particle wave functions $\Psi(\vec{r}_1, \vec{r}_2)$, the angular momentum acting on the first particle would be

  $$\langle \vec{r}_1, \vec{r}_2|\hat{\vec{L}}_1|\Psi\rangle =: \hat{\vec{L}}_1\Psi(\vec{r}_1, \vec{r}_2) = \frac{\hbar}{i}\vec{r}_1 \times \vec{\nabla}_{\vec{r}_1}\Psi(\vec{r}_1, \vec{r}_2)$$
  $$\langle \vec{r}_1, \vec{r}_2|\hat{L}_2|\Psi\rangle =: \hat{L}_2\Psi(\vec{r}_1, \vec{r}_2) = \frac{\hbar}{i}\vec{r}_2 \times \vec{\nabla}_{\vec{r}_2}\Psi(\vec{r}_1, \vec{r}_2)$$

  The total angular momentum is denoted as

  $$\hat{\vec{J}} \stackrel{\text{def}}{=} \hat{\vec{L}}_1 + \hat{\vec{L}}_2$$

  Its eigenvalues will be denoted by $j = \ell_J$ and $j_z = m_J$.

- Similarly we can consider a single particle with spin orbiting an atom. In this case, the total angular momentum has the form

  $$\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$$

  where $\hat{\vec{L}}$ is the orbital angular momentum and $\hat{\vec{S}}$ is the spin of the particle. In two-component spinor representation the total angular momentum has the form

  $$\begin{aligned}\hat{\vec{J}} &= \hat{\vec{r}} \times \hat{\vec{p}}\,\mathbf{1} + \hat{1}\;\frac{\hbar}{2}\vec{\boldsymbol{\sigma}} \\ &\widehat{=}\; \vec{r} \times \frac{\hbar}{i}\vec{\nabla}\begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix} + \frac{\hbar}{2}\vec{\sigma}\end{aligned}$$

  This is a difficult but very instructive example: We have operators, denoted by a hat, " ˆ ", that acts on states that correspond square integrable complex functions in the three-dimensional coordinate space and those that act on the two-dimensional spin states. The complication comes from the complex nature of the Hilbert space for that system. This Hilbert space is a product of a two-dimensional Hilbert space, describing the spin, and the Hilbert space of wave functions in three-dimensional coordinate space. A basis state in that space is

  $$|\vec{r}, \sigma\rangle$$

  where σ can have the values $\uparrow$ or $\downarrow$. The two dimensional matrices act on the second variable, whereas the differential operators act on the first variable.

When the electrons interact, each electron does not have a conserved angular momentum. The total angular momentum $\vec{J} = \vec{L}_1 + \vec{L}_2$ is still conserved. Hence we are looking for eigenstates of $\hat{\vec{J}}_z$ and $\hat{\vec{J}}^2$. Note that the eigenvalues for $\hat{\vec{J}}^2$ and $\hat{J}_z$ are denoted by $\hbar j(j+1)$ and $\hbar j_z$ respectively.

The question we address in the following is, how we transform between these two sets of eigenstates.

We start out with a basis of angular momentum eigenstates for each electron individually

$$\begin{aligned}
\hat{\vec{L}}_1^2|\ell_1, m_1, \ell_2, m_2\rangle &= |\ell_1, m_1, \ell_2, m_2\rangle\hbar^2\ell_1(\ell_1+1)\\
\hat{\vec{L}}_2^2|\ell_1, m_1, \ell_2, m_2\rangle &= |\ell_1, m_1, \ell_2, m_2\rangle\hbar^2\ell_2(\ell_2+1)\\
\hat{L}_{1,z}|\ell_1, m_1, \ell_2, m_2\rangle &= |\ell_1, m_1, \ell_2, m_2\rangle\hbar m_1\\
\hat{L}_{2,z}|\ell_1, m_1, \ell_2, m_2\rangle &= |\ell_1, m_1, \ell_2, m_2\rangle\hbar m_2\\
\hat{L}_{1,-}|\ell_1, m_1, \ell_2, m_2\rangle &= |\ell_1, m_1-1, \ell_2, m_2\rangle\hbar\sqrt{(\ell_1+m_1)(\ell_1-m_1+1)}\\
\hat{L}_{2,-}|\ell_1, m_1, \ell_2, m_2\rangle &= |\ell_1, m_1, \ell_2, m_2-1\rangle\hbar\sqrt{(\ell_2+m_2)(\ell_2-m_2+1)}\\
\hat{L}_{1,+}|\ell_1, m_1, \ell_2, m_2\rangle &= |\ell_1, m_1+1, \ell_2, m_2\rangle\hbar\sqrt{(\ell_1-m_1)(\ell_1-m_1+1)}\\
\hat{L}_{2,+}|\ell_1, m_1, \ell_2, m_2\rangle &= |\ell_1, m_1, \ell_2, m_2+1\rangle\hbar\sqrt{(\ell_2-m_2)(\ell_2-m_2+1)}
\end{aligned}$$

Let us express $\hat{\vec{J}}^2$ by the angular momenta of individual particles $\hat{\vec{L}}_1$ and $\hat{\vec{L}}_2$

$$\begin{aligned}
\hat{\vec{J}}^2 &= \hat{\vec{L}}_1^2 + 2\hat{\vec{L}}_1\hat{\vec{L}}_2 + \hat{\vec{L}}_2^2\\
&= \hat{\vec{L}}_1^2 + \hat{\vec{L}}_2^2 + 2\Big[\hat{L}_{1,x}\hat{L}_{2,x} + \hat{L}_{1,y}\hat{L}_{2,y} + \hat{L}_{1,z}\hat{L}_{2,z}\Big]\\
&= L_1^2 + L_2^2 + 2L_{1,z}L_{2,z} + \Big[L_{1,+}L_{2,-} + L_{1,-}L_{2,+}\Big]
\end{aligned}$$

Apparently the states are not eigenstates of the total angular momentum. To show this let us apply J_z and J^2 to the eigenstates of L_1 and L_2.

$$\begin{aligned}
J_z|\ell_1, m_1, \ell_2, m_2\rangle &= \hbar(m_1+m_2)|\ell_1, m_1, \ell_2, m_2\rangle\\
J^2|\ell_1, m_1, \ell_2, m_2\rangle &= \hbar^2[\ell_1(\ell_1+1) + \ell_2(\ell_2+1) + 2m_1m_2]|\ell_1, m_1, \ell_2, m_2\rangle\\
&\quad + |\ell_1, m_1+1, \ell_2, m_2-1\rangle\\
&\quad \times \hbar^2\sqrt{[\ell_1(\ell_1+1) - m_1(m_1+1)][\ell_2(\ell_2+1) - m_2(m_2-1)]}\\
&\quad + |\ell_1, m_1-1, \ell_2, m_2+1\rangle\\
&\quad \times \hbar^2\sqrt{[\ell_1(\ell_1+1) - m_1(m_1-1)][\ell_2(\ell_2+1) - m_2(m_2+1)]}
\end{aligned}$$

In order to find an eigenstate of J^2 we need to superimpose the states with a given ℓ_1, ℓ_2 and $m = m_1 + m_2$.

The new basis will have the eigenvalues

$$\begin{aligned}
J^2|\ell_1, \ell_2, \ell, m\rangle &= \hbar^2\ell(\ell+1)|\ell_1, \ell_2, \ell, m\rangle\\
J_z|\ell_1, \ell_2, \ell, m\rangle &= \hbar m|\ell_1, \ell_2, \ell, m\rangle\\
L_1^2|\ell_1, \ell_2, \ell, m\rangle &= \hbar^2\ell_1(\ell_1+1)|\ell_1, \ell_2, \ell, m\rangle\\
L_1^2|\ell_1, \ell_2, \ell, m\rangle &= \hbar^2\ell_1(\ell_1+1)|\ell_1, \ell_2, \ell, m\rangle
\end{aligned}$$

which is a definition of the basis functions $|\ell_1, \ell_2, \ell, m\rangle$.

The general form of the basis functions is

$$|\ell_1, \ell_2, \ell, m\rangle = \sum_{i=-\min(\ell_1,\ell_2)-m/2}^{\min(\ell_1,\ell_2)-m/2} |\ell_1, \frac{m}{2}+i, \ell_2, \frac{m}{2}-i\rangle c_i$$

where the coefficients c_i also depend on ℓ_1, ℓ_2, ℓ, m. The coefficients $c_i = \langle \ell_1, m_1, \ell_2, m_2 | \ell_1, \ell_2, \ell, m\rangle$ with $i = \frac{m_1 - m_2}{2}$, which also depend on ℓ_1, ℓ_2, m, ℓ are called **Clebsch-Gordan coefficients**, Wigner or vector coupling coefficients.

They are fairly complex so that we will not evaluate them here.

Editorial remark: Work on this section. Also include pictorial description of vector addition components as in Atkins

11.7 Products of spherical harmonics

Extremely useful is the product rule of spherical harmonics.

GAUNT COEFFICIENTS

The Gaunt coefficients[35] are defined as

$$G^{\ell'',m''}_{\ell,m,\ell',m'} \stackrel{\text{def}}{=} \int d\Omega\; Y^*_{\ell'',m''}(\vec{r}) Y_{\ell,m}(\vec{r}) Y_{\ell',m'}(\vec{r}) \tag{11.27}$$

The Gaunt coefficients can be used to express the product of two spherical harmonics by a sum of single spherical harmonics, called the **Gaunt series**.

$$Y_{\ell,m}(\vec{r}) Y_{\ell',m'}(\vec{r}) = \sum_{\ell'',m''} G^{\ell'',m''}_{\ell,m,\ell',m'} Y_{\ell'',m''}(\vec{r})$$

Note that the definition of Gaunt coefficients is not uniform, and I have seen definitions with an additional factor of $\sqrt{4\pi}$ (Cohen Tannoudji). In order to confirm numerically of the definition agrees with the one given above, simply test the coefficients $G^{0,0}_{\ell,m,\ell,m} = \frac{1}{\sqrt{4\pi}}$. This rule can be derived, from knowing that $Y_{\ell,m} = \frac{1}{\sqrt{4\pi}}$.

The product rule is for example needed to express the particle density of an atom in spherical harmonics, if the wave function are given as radial function times spherical harmonics. If $\Psi_n(\vec{r}) = R_n(|\vec{r}|) Y_{\ell,m}(\vec{r})$, we obtain

$$\rho(\vec{r}) = q\Psi^*(\vec{r})\Psi(\vec{r}) = \sum_{\ell'',m''} G^{\ell'',m''}_{\ell,-m,\ell,m} |R(|\vec{r}|)|^2\, Y_{\ell'',m''}(\vec{r})$$

Here, we used the property that $Y^*_{\ell,m}(\vec{r}) = Y_{\ell,-m}(\vec{r})$, which is obtained from the definition Eq. 11.23 with the help of Eq. 11.25.

The Gaunt coefficients are closely related to the Clebsch-Gordan coefficients:

$$G^{\ell,m}_{\ell_1,m_1,\ell_2,m_2} = (-1)^{m_2} \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)}{4\pi(2\ell+1)}} \langle \ell_1, 0, \ell_2, 0 | \ell_1, \ell_2, \ell, 0\rangle \langle \ell_1, m_1, \ell_2, -m_2 | \ell_1, \ell_2, \ell, m\rangle$$

In the following we sketch how these coefficients are derived: We start out with a two particle wave function, which depends on the positions $\vec{r}_1$ and $\vec{r}_2$ of the two particles.

$$\langle r_1, r_2 | \ell_1, m_1, \ell_2, m_2\rangle = Y_{\ell_1,m_1}(\vec{r}_1) Y_{\ell_2,m_2}(\vec{r}_2)$$
$$\langle r_1, r_2 | \ell_1, \ell_2, \ell, m\rangle = \Phi(\vec{r}_1, \vec{r}_2)$$

We can relate Φ to the products of two wave functions using the Clebsch Gordan coefficients.

$$\langle \vec{r}_1, \vec{r}_2 | \ell_1, m_1, \ell_2, m_2\rangle = \sum_{\ell,m} \langle \vec{r}_1, \vec{r}_2 | \ell_1, \ell_2, \ell, m\rangle \langle \ell_1, \ell_2, \ell, m | \ell_1, m_1, \ell_2, m_2\rangle$$
$$Y_{\ell_1,m_1}(\vec{r}_1) Y_{\ell_2,m_2}(\vec{r}_2) = \sum_{\ell,m} \Phi_{\ell_1,\ell_2,\ell,m}(\vec{r}_1, \vec{r}_2) \langle \ell_1, \ell_2, \ell, m | \ell_1, m_1, \ell_2, m_2\rangle$$

and therefore also

$$Y_{\ell_1,m_1}(\vec{r})Y_{\ell_2,m_2}(\vec{r}) = \sum_{\ell,m} \Phi_{\ell_1,\ell_2,\ell,m}(\vec{r},\vec{r})\langle \ell_1,\ell_2,\ell,m|\ell_1,m_1,\ell_2,m_2\rangle$$

It can be shown that $\Phi(\vec{r},\vec{r})$ is proportional to a spherical harmonic itself. Once that is shown, we only need to work out the normalization constants.

Let us show that $F_{\ell_1,\ell_2,\ell,m}(\vec{r}) = \Phi_{\ell_1,\ell_2,\ell,m}(\vec{r},\vec{r})$ is a spherical harmonic. Therefore, we apply an angular momentum operator to F: Using product and chain rule we relate the derivatives of F. We can consider $\vec{r}_1$ and $\vec{r}_2$ to be functions of r namely $\vec{r}_1(\vec{r}) = \vec{r}$ and $\vec{r}_2(\vec{r}) = \vec{r}$.

$$\begin{aligned}
\vec{r}_1(\vec{r}) &= \vec{r} \\
\vec{r}_2(\vec{r}) &= \vec{r} \\
\vec{r}\times\nabla F(\vec{r}) &= \vec{r}\times\nabla\Phi(\vec{r}_1(\vec{r}),\vec{r}_2(\vec{r})) \\
&= [\vec{r}\times\nabla_1 + \vec{r}\times\nabla_2]\Phi(\vec{r}_1(\vec{r}),\vec{r}_2(\vec{r})) \\
&= [\vec{r}_1\times\nabla_1 + \vec{r}_2\times\nabla_2]\Phi(\vec{r}_1(\vec{r}),\vec{r}_2(\vec{r})) \\
LF(\vec{r}) &= [\vec{L}_1 + \vec{L}_2]\Phi(\vec{r},\vec{r})
\end{aligned}$$

where $\vec{L}_1$ acts only on the first argument and $\vec{L}_2$ on the second.

Because $\Phi(\vec{r}_1,\vec{r}_2)$ is an eigenstate of angular momentum $(\vec{L}_1+\vec{L}_2)\Phi(\vec{r}_1,\vec{r}_2) = \Phi(\vec{r}_1,\vec{r}_2)\hbar^2\ell(\ell+1)$ and $(\vec{L}_{1,z}+\vec{L}_{2,z})\Phi(\vec{r}_1,\vec{r}_2) = \Phi(\vec{r}_1,\vec{r}_2)\hbar m$ we find that also $F(r)$ is an eigenstate of angular momentum L_z and L^2,

$$\begin{aligned}
L_z|F\rangle &= \hbar m|F\rangle \\
L^2|F\rangle &= \hbar^2\ell(\ell+1)|F\rangle
\end{aligned}$$

which implies that they are related to spherical harmonics.

The derivation of the remaining coefficients can be found in Cohen-Tannoudji[3].

11.8 Recommended exercises

1. Rotator: Exercise 16.5 on p. 246. (Not ready)

Chapter 12

Atoms

12.1 Radial Schrödinger equation

Let us consider a particle in a spherically symmetric potential

$$\left[\frac{\hat{\vec{p}}^2}{2m}+V(|\hat{\vec{r}}|)\right]|\psi\rangle=|\psi\rangle E$$

Editor: Include hats to distinguish numbers from operators

We introduce the **radial momentum** $\hat{p}_r=\frac{1}{|r|}(\hat{\vec{r}}\hat{\vec{r}})$ and obtain an expression that contains radial and angular momenta separately.

$$\left[\frac{1}{2m}\left(\frac{1}{\hat{r}}\hat{p}_r^2\hat{r}+\frac{1}{\hat{r}^2}\hat{L}^2\right)+V(\hat{r})\right]|\psi\rangle=|\psi\rangle E$$

where $r=|\vec{r}|$. Note, that[8]

$$\frac{1}{r}p_r^2r=\frac{1}{r^2}p_r\left(r^2p_r\right)$$

Here we show that $\vec{p}^2$ can be written as

$$\vec{p}^2 = \frac{1}{r}p_r^2 r + \frac{\vec{L}^2}{r^2}$$

PROOF: We start with squaring one component of the angular momentum

$$\begin{aligned}
L_x^2 &= (yp_z - zp_y)^2 = yp_z yp_z - yp_z zp_y - zp_y yp_z + zp_y zp_y \\
&= y^2 p_z^2 - y[p_z, z]p_y - yzp_z p_y - z[p_y, y]p_z - zyp_y p_z + z^2 p_y^2 \\
&= (y^2 p_z^2 + z^2 p_y^2) - 2(yp_y)(zp_z) - \frac{\hbar}{i}(yp_y + zp_z)
\end{aligned}$$

with cyclic permutation we find

$$\begin{aligned}
\vec{L}^2 &= (y^2 p_z^2 + z^2 p_y^2) - 2(yp_y)(zp_z) - \frac{\hbar}{i}(yp_y + zp_z) \\
&+ (z^2 p_x^2 + x^2 p_z^2) - 2(zp_z)(xp_x) - \frac{\hbar}{i}(zp_z + xp_x) \\
&+ (x^2 p_y^2 + y^2 p_x^2) - 2(xp_x)(yp_y) - \frac{\hbar}{i}(xp_x + yp_y) \\
&= \Big(x^2(p_y^2 + p_z^2) + y^2(p_x^2 + p_z^2) + z^2(p_x^2 + p_y^2)\Big) \\
&- \Big(xp_x(yp_y + zp_z) + yp_y(xp_x + zp_z) + zp_z(xp_x + yp_y)\Big) \\
&- \frac{2\hbar}{i}\Big(xp_x + yp_y + zp_z\Big)
\end{aligned}$$

Now we add the identity $x^2 p_x^2 - (xp_x)^2 + \frac{\hbar}{i}xp_x = 0$, and those obtained by exchanging x by y and z to the expression above

$$\begin{aligned}
\vec{L}^2 &= \vec{r}^2\vec{p}^2 - (\vec{r}\vec{p})^2 - \frac{\hbar}{i}\vec{r}\vec{p} \\
&= \vec{r}^2\vec{p}^2 - (rp_r)\Big(rp_r - \frac{\hbar}{i}\Big) \\
&= r^2\vec{p}^2 - (rp_r)(p_r r) \\
\vec{p}^2 &= \frac{1}{r}p_r^2 r + \frac{1}{r^2}L^2
\end{aligned}$$

where we used $\vec{r}\vec{p} = rp_r$ and $[p_r, r] = \frac{\hbar}{i}$. *q.e.d.*

Now we use as Ansatz the eigenstates of the angular momentum

$$\psi(\vec{r}) = R_\ell(|\vec{r}|)Y_{\ell,m}(\vec{r})$$

By insertion we find[1]

$$\begin{aligned}
0 &= \Big[\frac{1}{2m}\Big(\frac{1}{r}p_r^2 r + \frac{1}{r^2}L^2\Big) + V(r) - E\Big]R_\ell(r)Y_{\ell,m}(\vec{r}) \\
&= Y_{\ell,m}(\vec{r})\Big[\frac{1}{2m}\Big(\frac{1}{r}p_r^2 r + \frac{\hbar^2\ell(\ell+1)}{r^2}\Big) + V(r) - E\Big]R_\ell(r)
\end{aligned}$$

We can divide by the spherical harmonics and obtain an one-dimensional differential equation for the radial component of the wave function.

We can rewrite[2] the radial Schrödinger equation in the form

$$\Big[\frac{-\hbar^2}{2m}\partial_r^2 + \frac{\hbar^2\ell(\ell+1)}{2mr^2} + V(r) - E\Big]rR(r) = 0$$

[1] We used here that the radial momentum commutates with the spherical harmonics, which is obvious, as the spherical harmonics do not depend on the radial coordinate.

[2] Multiply from the left with r.

Thus the radial Schrödinger equation has the form of a regular one-dimensional Schrödinger equation with an angular dependent potential, that also includes a potential which represents the centrifugal force. It is as if a one-dimensional particle would move in a potential $V_\ell(r) = \frac{\hbar^2\ell(\ell+1)}{2m}\frac{1}{r^2} + V(r)$. For an electron in that experiences the electrostatic attraction to the nucleus $V(r) = -\frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r}$. The ℓ-dependent potentials for the hydrogen atom are shown in Fig. 12.1.

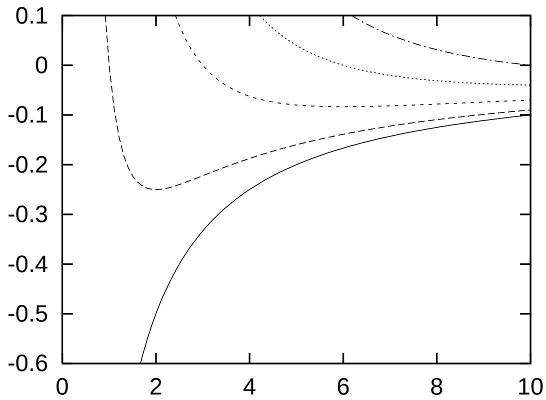


Fig. 12.1: Radial potential of the hydrogen atom (Z=1) including the centrifugal part for $\ell = 0, 1, 2, 3, 4$. The minima are at $r_{min} = \frac{4\pi\epsilon_0\hbar^2}{me^2}\frac{\ell(\ell+1)}{Z}$ and the well depth is $V_{min} = -\frac{1}{2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \frac{m}{\hbar^2}\frac{Z^2}{2\ell(\ell+1)}$.

Formal relation of the radial Schrödinger equation to Newton's equation of motion

The radial Schrödinger equation for $rR(r)$ is an ordinary differential equation, and can also be solved with the methods developed for the latter. It may be instructive to show the relation to Newton's equations of motion. Because we may have a good imagination on the behavior of classical particles compared to that of quantum mechanical wave functions.

Consider the Newtons equation of motion for a harmonic oscillator with a time-dependent force constant.

$$m\partial_t^2 x(t) = -c(t)x(t)$$

We do now the following transformations in order to convert Newtons equation of motion into the radial Schrödinger equation.

- The time is converted into the radial distance, i.e. $t \to r$
- the position $x(t)$ is converted into $rR(r)$
- the mass of the harmonic oscillator is replaced by $\frac{\hbar^2}{2m}$
- the force constant $c(t)$ is converted into the radial kinetic energy $E - \left(V(r) + \frac{\hbar^2\ell(\ell+1)}{2mr^2}\right)$

$$\begin{aligned} \left[-m\partial_t^2 - c(t)\right]x(t) &= 0 \\ \to \left[-\frac{\hbar^2}{2m}\partial_r^2 + V(r) - E\right] rR(r) &= 0 \end{aligned}$$

Boundary conditions at the origin

The main difference of the radial Schrödinger equation to the one-dimensional Schrödinger equation are the boundary conditions at the origin $r = 0$. We proceed here with an expansion of R in powers of r, that is

$$R(r) = \sum_j a_j r^j$$

We maintain only the leading order in r of the potential, that is $V(r) = Cr^\alpha\big(1 + O(r)\big)$.

$$\begin{aligned}
0 &= \Big[\frac{-\hbar^2}{2m}\partial_r^2 + \frac{\hbar^2\ell(\ell+1)}{2mr^2} + Cr^\alpha - E\Big]\sum_j a_j r^{j+1} \\
&= \sum_j \Big[\frac{-\hbar^2}{2m}j(j+1)a_j r^{j-1} + \frac{\hbar^2\ell(\ell+1)}{2m}a_j r^{j-1} + Ca_j r^{\alpha+j+1} - Ea_j r^{j+1}\Big] \\
&= \sum_j r^{j-1}\Big[\frac{-\hbar^2}{2m}\Big(j(j+1) - \ell(\ell+1)\Big)a_j + Ca_{j-\alpha-2} - Ea_{j-2}\Big]
\end{aligned}$$

which results in a recursion relation for the coefficients

$$\begin{aligned}
0 &= \frac{-\hbar^2}{2m}\Big(j(j+1) - \ell(\ell+1)\Big)a_j + Ca_{j-\alpha-2} - Ea_{j-2} \\
&= (j-\ell)(j+\ell+1)a_j - \frac{2mC}{\hbar^2}a_{j-\alpha-2} + \frac{2mE}{\hbar^2}a_{j-2} \\
a_j &= \frac{1}{(j-\ell)(j+\ell+1)}\Big(\frac{2mC}{\hbar^2}a_{j-\alpha-2} - \frac{2mE}{\hbar^2}a_{j-2}\Big) \qquad (12.1)
\end{aligned}$$

Let us consider the case where

- $\alpha = -1$. This case is important for atoms, because the electrostatic potential of an electron that experiences the Coulomb attraction of the nucleus is of this form $V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}$. We need to consider a solution for which $a_{j<j_0} = 0$. In Eq. 12.1, a_{j_0} can only be nonzero, if the denominator vanishes, which determines the possible values of j_0.

  $$(j_0 - \ell)(j_0 + \ell + 1)a_{j_0} = 0 \Rightarrow j_0 = \ell \text{ or } \; j_0 = -\ell - 1$$

  The solution $j_0 = -\ell - 1$ can be discarded because it will be irregular at the origin and not normalizable. The coefficient a_{j_0} is arbitrary and will be determined by the normalization condition.

  In order to obtain the next term in the expansion, we insert this result into Eq. 12.1 for $j = j_0 + 1 = \ell + 1$

  $$a_{\ell+1} = \frac{mC}{\hbar^2}\frac{1}{\ell+1}a_\ell$$

  Thus the radial wave function behaves near the origin as

  $$R_r(r) = A\Big[1 + \frac{mC}{\hbar^2(\ell+1)}r + O(r^2)\Big]r^\ell$$

  and is not differentiable for $\ell = 0$, which is a consequence of the Coulomb singularity. This kink is not energy dependent, while the higher order terms would depend explicitly on energy.

  For the "hydrogen atom" with atomic number Z we obtain $C = -Ze^2/(4\pi\epsilon_0)$ and a recursion relation with $k = j - \ell$

  $$\begin{aligned}
  a_j &= \frac{1}{(j-\ell)(j+\ell+1)}\Big(-\frac{2mZe^2}{4\pi\epsilon_0\hbar^2}a_{j-1} + \frac{2mE}{\hbar^2}a_{j-2}\Big) = 0 \\
  a_{k<0} &= 0
  \end{aligned}$$

so that

$$R_r(r) = a_0\Big[1 - \frac{mZe^2}{4\pi\epsilon_0\hbar^2(\ell+1)}r + O(r^2)\Big]r^\ell$$

Note, that the kink is independent of the energy.

- $\alpha = 0$ which describes a continuous potential. The term proportional $r^{\ell+1}$ vanishes, and the second radial function behaves like

$$R_r(r) = A\Big[1 + O(r^2)\Big]r^\ell$$

The recursion relation Eq. 12.1 for the coefficients could be used to determine a Taylor expansion of $R(r)$ to arbitrary high order. More accurate results are obtained, when one uses the Taylor expansion to obtain value and derivative at a point close to the nucleus and to use these values as initial conditions for a numerical solution of the Schrödinger equation.

12.2 The Hydrogen Atom

The potential for the hydrogen atom is

$$V(\vec{r}) = -\frac{Ze^2}{4\pi\epsilon_0|r|}$$

where Z is the atomic number, e is the elementary charge, and $\vec{r}$ is the position relative to the nuclear site. The radial potentials including the centrifugal term are shown in Fig. 12.1.

$$\Big[\frac{-\hbar^2}{2m_e}\partial_r^2 + \frac{\hbar^2\ell(\ell+1)}{2m_e r^2} - \frac{Ze^2}{4\pi\epsilon_0 r} - E\Big]rR(r) = 0$$

Hartree atomic units

Let us simplify the equation by introducing dimensionless quantities. We multiply with $m_e/\hbar^2$ and introduce a length unit and an energy unit so that the equation is dimensionally consistent.

$$\begin{aligned}
\Big[-\partial_r^2 + \frac{\ell(\ell+1)}{r^2} + \frac{2m_eZe^2}{4\pi\epsilon_0\hbar^2}\frac{1}{r} - \frac{2m_eE}{\hbar^2}\Big]rR(r) &= 0\\
\Big[-\frac{1}{2}\partial_r^2 + \frac{\ell(\ell+1)}{2r^2} + \frac{Z}{a_0 r} - \frac{E}{E_0a_0^2}\Big]rR(r) &= 0\\
\Big[-\frac{1}{2}\partial_y^2 + \frac{\ell(\ell+1)}{2y^2} + \frac{Z}{y} - \frac{E}{E_0}\Big]yR(ya_0) &= 0 \qquad (12.2)
\end{aligned}$$

Thus we have defined a length scale, the Bohr radius a_0, and an energy scale, the Hartree $E_0 = H$.

Definition 12.1 *BOHR RADIUS*

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} \qquad (12.3)$$

The Bohr radius is the length unit of the Hartree atomic unit system. It corresponds to the radius of a classical electron circling a proton at the ground state energy of the quantum mechanical hydrogen atom.

Definition 12.2 *HARTREE*

$$H = \frac{\hbar^2}{m_e a_0^2} = \frac{m_e e^4}{(4\pi\epsilon_0\hbar)^2} \tag{12.4}$$

The Hartree is the energy unit of the Hartree atomic unit system. One Hartree is twice the binding energy of an electron in the hydrogen atom.

In this way we can define a new system of units, the so-called **Hartree atomic units**.

Definition 12.3 *HARTREE ATOMIC UNITS*
The Hartree atomic units are defined by

$$\hbar = m_e = e = 4\pi\epsilon_0 = 1$$

where m_e is the electron mass, e is the elementary charge, and ϵ_0 is the dielectric constant of the vacuum.
The Hartree atomic unit system is commonly used for atomic and molecular physics. Unlike the official SI unit system, Hartree atomic units are defined entirely by measurable fundamental quantities.

- *the length unit of the Hartree atomic unit system is the Bohr radius, which corresponds to the radius of the classical orbit of an electron in the hydrogen atom.*
- *the energy unit is the Hartree, which is twice the binding energy of a hydrogen atom.*

Instead of using the symbols a_0, H, etc. one often uses a.u. which stands for "atomic unit" in general. Note that there are two atomic unit systems. The so-called **Rydberg atomic unit system** *is different. This leads to confusion, because often it is not spelled out, which of the two atomic unit systems is meant.*

In Hartree atomic units, the Schrödinger equation of an electron in the Coulomb potential of the nucleus has the form

$$\Big(-\frac{1}{2}\nabla^2 - \frac{Z}{r} - E\Big)|\Psi\rangle = 0$$

The values of the fundamental constants and the conversion factors can be found in the book "Quantities, Units and Symbols in Physical Chemistry" [36] or on the web[37, 38].

Bound wave functions

Before we solve Eq. 12.2 let us investigate its qualitative behavior, which will guide us to useful ansatz for the solution. We can simplify the differential equation, if we consider the two limits $r \to 0$ and $r \to \infty$

- For $r \to 0$ we can use the Power series expansion, which tells us that the wave function at the origin is proportional to r^ℓ.
- For $r \to \infty$ we can ignore the potential and the centrifugal term. Thus the differential equation becomes independent of Z and ℓ. We obtain with $y = r/a0$

  $$\Big[-\frac{1}{2}\partial_y^2 - \frac{E}{H}\Big] y R(y a_0) \approx 0$$

  which can easily be solved. The solution has the form

  $$R(y a_0) \sim \frac{1}{y} \mathrm{e}^{-\sqrt{2\frac{E}{H}}\, y}$$

The wave function decays exponentially with the decay governed by the energy. For bound states the wave functions decay exponentially, while they oscillate for positive energies. Here we only consider the bound states.

These considerations suggest an Ansatz of the form

$$R(ya_0) = F(y)y^{\ell}\mathrm{e}^{-\lambda x}$$

where

$$\lambda \stackrel{\text{def}}{=} \sqrt{2\frac{E}{H}} \tag{12.5}$$

The goal is now to determine the function $F(y)$

We insert this Ansatz into the radial Schrödinger equation. Eq. 12.2:

$$\Big[-\frac{1}{2}\partial_y^2 + \frac{\ell(\ell+1)}{2y^2} - \frac{Z}{y} - \frac{E}{H}\Big] F(y)y^{\ell+1}\mathrm{e}^{-\lambda y} = 0 \ .$$

We use

$$\begin{aligned}
\partial_y y^{\ell+1}\mathrm{e}^{-\lambda y} &= y^{\ell+1}\mathrm{e}^{-\lambda y}[\frac{\ell+1}{y} - \lambda + \partial_y] \\
\Rightarrow \partial_y^2 y^{\ell+1}\mathrm{e}^{-\lambda y} &= y^{\ell+1}\mathrm{e}^{-\lambda y}[\frac{\ell+1}{y} - \lambda + \partial_y][\frac{\ell+1}{y} - \lambda + \partial_y] \\
&= y^{\ell+1}\mathrm{e}^{-\lambda y}\left[\left(\frac{\ell+1}{y} - \lambda\right)^2 + \partial_y\left(\frac{\ell+1}{y} - \lambda\right) + \left(\frac{\ell+1}{y} - \lambda\right)^2 \partial_y + \partial_y^2\right] \\
&= y^{\ell+1}\mathrm{e}^{-\lambda y}\left[(\frac{\ell+1}{y} - \lambda)^2 - \frac{\ell+1}{y^2} + 2[\frac{\ell+1}{y} - \lambda]\partial_y + \partial_y^2\right] \\
&= y^{\ell+1}\mathrm{e}^{-\lambda y}\left[\partial_y^2 + 2[\frac{\ell+1}{y} - \lambda]\partial_y + \frac{\ell(\ell+1)}{y^2} - 2\lambda\frac{\ell+1}{y} + \lambda^2\right]
\end{aligned}$$

and insert into the radial Schrödinger equation Eq. 12.2

$$\Big[\partial_y^2 + 2[\frac{\ell+1}{y} - \lambda]\partial_y - \frac{2}{y}\Big(\lambda(\ell+1) - Z\Big) + \underbrace{\Big(\frac{2E}{H} + \lambda^2\Big)}_{=0}\Big] F(y) = 0$$

$$\Rightarrow \Big[y\partial_y^2 + 2[\ell+1-\lambda y]\partial_y - 2\Big(\lambda(\ell+1) - Z\Big)\Big] F(y) = 0 \tag{12.6}$$

This differential equation is a special case of the general **associated Laguerre differential equation**.

$$\left(x\partial_x^2 + (k+1-x)\partial_x + n\right) L_n^{(k)}(x) = 0$$

where k, n are integers. The solutions are the so-called **associated Laguerre Polynomials**

$$L_n^{(k)}(x) = \frac{\mathrm{e}^x x^{-k}}{n!}\frac{d^n}{dx^n}\mathrm{e}^{-x}x^{n+k}$$

This can be used if one find the corresponding tables. A useful source is the Book of Abramowitz and Stegun[39].

However, here we want to investigate how such a problem can be solved directly: We make a power-series Ansatz for $F(y) = \sum_j a_j y^j$ in Eq. 12.6.

$$\sum_j \Big[j(j-1)a_j y^{j-1} + 2(\ell+1)ja_j y^{j-1} - 2\lambda j a_j y^j - 2\Big(\lambda(\ell+1) - Z\Big)a_j y^j\Big] = 0$$

$$\sum_j [j(j-1) + 2(\ell+1)j]\, a_j y^{j-1} - \sum_j \Big[2\lambda j + \Big(\lambda(\ell+1) - Z\Big)\Big]\, a_j y^j = 0$$

$$\sum_j \left\{ [(j+1)j + 2(\ell+1)(j+1)]\, a_{j+1} - \sum_j \Big[2\lambda j + \Big(\lambda(\ell+1) - Z\Big)\Big]\, a_j \right\} y^j = 0$$

Because the identity holds for all values of y, it must also hold in every power of y.[3] If we collect all terms that are proportional to a certain power of y, namely y^{j-1}, we obtain

$$
\begin{aligned}
&j(j+1)a_{j+1}+2(\ell+1)(j+1)a_{j+1}-2\lambda j a_j-2\Big(\lambda(\ell+1)-Z\Big)a_j=0\\
\Rightarrow\;&\Big(j(j+1)+2(\ell+1)(j+1)\Big)a_{j+1}-2\Big(\lambda j+\lambda(\ell+1)-Z\Big)a_j=0,
\end{aligned}
$$

which provides us with a recursion relation for the coefficients a_j

$$
a_{j+1}=2\frac{\lambda(j+\ell+1)-Z}{(j+1)(j+2\ell+2)}a_j \tag{12.7}
$$

The power series can start following that index, where the denominator vanishes that is with $j=0$ or $j=-(2\ell+1)$. The second choice leads to the **irregular solution**. The irregular solution has a singularity at the origin. The resulting wave function would obey the Schrödinger equation everywhere, except at the origin.Therefore, the irregular solution must be excluded. Thus the power series starts with $j_{min}=0$, which leads to the **regular solution**.

The power series also has an upper end, where the numerator of Eq. 12.7 vanishes: The power series expansion finishes, if there is an integer j_{max} so that

$$
0=\lambda(j_{max}+\ell+1)-Z \qquad\Rightarrow\qquad j_{max}=-\ell-1+\frac{Z}{\lambda}
$$

j_{max} must itself be a positive integer, because otherwise the power series does not terminate. (Some further calculations are required to show that such solutions lead to exponentially increasing functions that cannot be normalized.[4])

The requirement that j_{max} is an integer, determines the allowed values for λ. For j_{max} we obtain

$$
\lambda_{j_{max}}=\frac{Z}{j_{max}+\ell+1} \tag{12.8}
$$

We introduce a new variable the so-called **principal quantum number**

$$
n=j_{max}+\ell+1 \tag{12.9}
$$

Now we take the definition of λ, Eq. 12.5, resolve it for the energy,

$$
E=\frac{1}{2}\lambda^2\;H
$$

and insert Eq. 12.8 with j_{max} expressed by the main quantum number as defined in Eq. 12.9. Remember that H is the Hartree energy unit defined in Eq. 12.4

Thus, we obtain an expression for the allowed energy eigenvalues and thus the eigenvalue spectrum of the bound states of the hydrogen atom.

[3]These becomes obvious by forming higher derivatives of the expression, which must all vanish, because the expression is zero for all values of y.

[4]Power series expansions that do not truncate will fulfill an approximate recursion relation

$$
\begin{aligned}
a_{j+1}&\approx 2\frac{\lambda-Z}{(j+1)}a_j\\
a_j&\approx \frac{1}{j!}2^j(\lambda-Z)^j\\
F(x\to\infty)&\to \mathrm{e}^{2(\lambda-Z)x}
\end{aligned}
$$

Thus $F(x)$ will diverge exponentially, which violates the boundary conditions at infinity.**Editor: For $\lambda<Z$ the function $F(x)$ does not converge! Check if those values are allowed and what is their meaning!**

EIGENVALUE SPECTRUM OF THE HYDROGEN ATOM

$$E_{n,\ell,m} = -\frac{Z^2}{2n^2}\,H = -\frac{Z^2}{2n^2}\frac{m_e e^4}{(4\pi\epsilon_0\hbar)^2} \tag{12.10}$$

for $n = j_{max} + \ell + 1 = 1, 2, 3, \ldots$.
States for different ℓ and m quantum numbers are degenerate. n is the principal quantum number, ℓ the angular momentum quantum number, and m is called the magnetic quantum number.
Only states with $\ell \geq n - 1$ contribute to one multiplet of degenerate states.

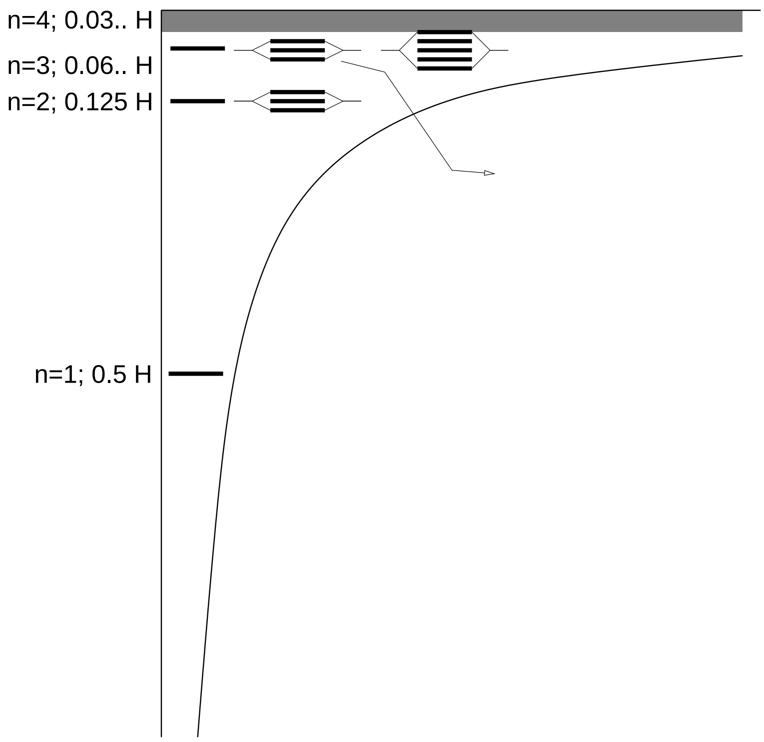


Fig. 12.2: Electron levels of the hydrogen atom.

We can now write down the recursion relation for the coefficients depending on n as

$$\lambda_n = \sqrt{\frac{-2E_n}{H}} = \frac{Z}{n}$$
$$a_{j+1} = \frac{2Z}{n}\frac{j+\ell+1-n}{(j+1)(j+2\ell+2)}a_j$$
$$R_{n,\ell}(r) = \sum_j b_j\left(\frac{Zr}{na_0}\right)^{j+\ell}\mathrm{e}^{-Zr/(na_0)}$$
$$b_{j+1} = 2\frac{j+\ell+1-n}{(j+1)(j+2\ell+2)}b_j$$

Note that these wave functions are not yet normalized. The polynomials appearing in this equation are, up to a normalization constant, the associated Laguerre polynomials.

$$R(r) = L^{2\ell+1}_{n+\ell}(2\frac{Zr}{na_0})r^\ell\mathrm{e}^{-\frac{Zr}{na_0}}$$

n	l	$E[H]$	R(r)
1	0	$-\frac{Z^2}{2}$	$2(a_0)^{-\frac{3}{2}}\mathrm{e}^{-r/a_0}$
2	0	$-\frac{Z^2}{8}$	$2(2a_0)^{-\frac{3}{2}}\left(1-\frac{r}{2a_0}\right)\mathrm{e}^{-r/(2a_0)}$
2	1	$-\frac{Z^2}{8}$	$2(2a_0)^{-\frac{3}{2}}\frac{r}{a_0}\mathrm{e}^{-r/(2a_0)}$

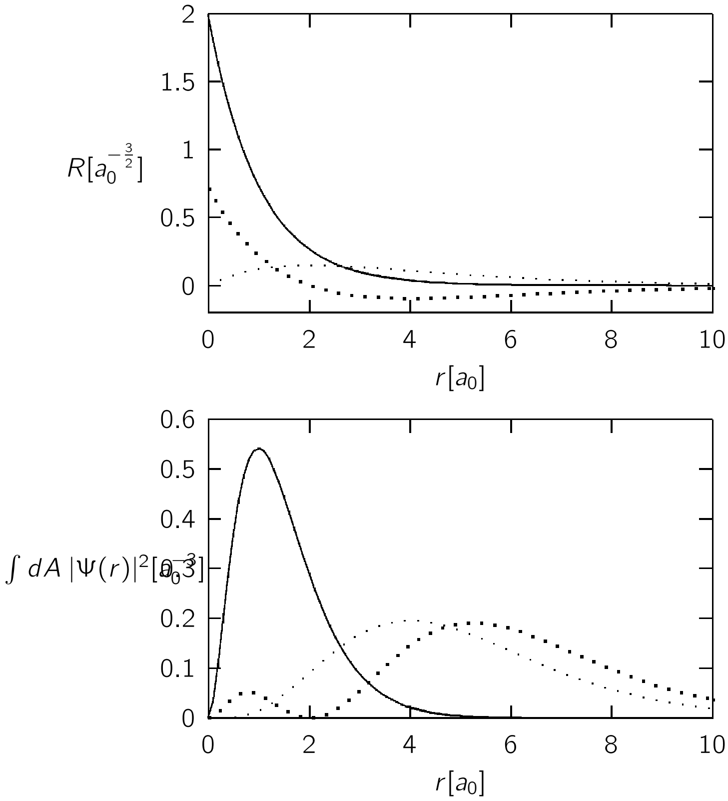


Fig. 12.3: Radial parts of the 1s,2s and 2p wave function of the hydrogen atom. Formulas taken from [4]. These functions are normalized. The results need to be checked.**Editor: check formulas and refine figures**

Two observations

- The number of node planes of the wave function increases with n. Angular node planes must be counted as well.

- wave functions and energies for general atomic number can be obtained from the hydrogen atom by a simple scaling law.

Periodic Table

There are states with angular momentum greater than zero are degenerate with those with $\ell = 0$. This explains already a large part of the structure of the periodic table.

H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	Ac															

- the lowest state in singly degenerate (or doubly degenerate if we include spin degeneracy). Thus we can fill this state with one or two electrons resulting in H and He.
- The next shell contains one state with $\ell = 0$ and three states with $\ell = 1$. The ones with only s-electrons ($\ell = 0$) are Li and Be. Alkali and Earth alkali metals. Then come the ones with partially filled p-shell ($\ell = 1$), B,C,N,O,F. The last atom in this row is Ne a noble gas that does not form any chemical bonds, because it fulfills the **octet rule**. The octet rule is fulfilled if the (s and p like valence wave functions) are occupied with just eight electrons.
- Electronegativity: The electrons in the first groups (columns in the periodic table) are loosely bound, while those in the last group have the most strongly bound electrons, because the electron energies decrease with increasing atomic number. Atoms with low lying electron levels are called electronegative. Noble gas atoms, in the rightmost column of the periodic table, are inert, because on the one hand they hold on to their electrons most strongly. On the other hand they cannot accept any more electrons, because those would need to occupy the next shell having much higher energies. If we combine atoms from an early group and one of a late group, we form ionic materials and electrons are transferred so that the atom from an early group is positive (cation) and the one from a late group is negative (anion) as in NaCl. Anions can accept electrons until their shell is filled, while cations donate electrons until their shell is empty. Thus we can estimate the composition of most oxides: Na_2O, MgO, Al_2O_3, SiO_2 etc.

  Because electrons in the early groups are loosely bound, they often form metals such as Na, Mg, where the electrons are free to wander around. In contrast, towards the right of the periodic table they form covalently bound materials such as F_2, O_2, N_2.
- Due to Coulomb repulsion between electrons the ℓ degeneracy is lifted in real atoms, so that the orbitals with lower angular momentum are lower in energy than those with higher angular momentum. The reason is that the valence electrons experience a lower effective atomic number than the core electrons. The electrostatic potential of the nucleus is shielded by the core electrons. This makes the core states more compact. The valence states with low angular momentum can therefore come closer to the nucleus as well, because they are kept at larger radius mostly by the Pauli repulsion from the core electrons. As a consequence the s-electrons are occupied before p-electrons are occupied.

  This effect also explains an anomaly in the periodic table, namely that the third period contains only s and p electrons, while the 3d ($n = 3, \ell = 2$) electrons are occupied only after the 4s ($n = 4, \ell = 0$) electrons have been filled.

Chapter 13

Approximation techniques

In most cases quantum mechanical problems cannot be solved exactly. Therefore, for real-world problems, we usually need to resort to approximations or numerical methods, or a combination of both. There is no limit to the number of approximation techniques one can use for certain problems. However, there is a set of techniques, that already get us a long way, and the ideas behind these approaches can be used to use new approximations for a particular problem.

13.1 Perturbation theory

Perturbation theory can be used to determine the properties of a complex system if a similar system that is much simpler can explicitly be solved. For example the complex system may have an approximate symmetry. We can solve the system that has the symmetry exactly fulfilled, which is simpler, and then work out the differences of the real system to our model system.

Perturbation theory is not only an approximation method but in many cases it provides information more directly than a brute force approach. This is the case when we are interested in the linear response of a quantity to a perturbation. For example we would like the splitting of energy levels of an atom in an external field, because the splitting may experimentally be accessible. There is not point to evaluate all the energy values, when it is possible to calculate the splitting directly.

The first order perturbation is also what is needed in **linear response theory.** For example we would like to evaluate the dielectric constant of a gas. The dielectric constant is related to the linear response[1] of the polarization $\vec{P}$ to an electric field. This response is called the electric susceptibility.

Let us consider a system that can be described by an unperturbed Hamiltonian $\hat{H}_0$ and a small perturbation $\hat{W}$.[2] We assume that we can solve the Schrödinger equation for the unperturbed system, and we want to use that knowledge to obtain an estimate for the perturbed system. The full Hamiltonian $\hat{H}$ is given as

$$\hat{H} = \hat{H}_0 + \hat{W}$$

If the perturbation is small, we need not solve the entire problem, but we can instead use perturbation theory. The basic idea of the procedure is as follows: We introduce a scale parameter λ for the perturbation so that

$$\hat{H}(\lambda) = \hat{H}_0 + \lambda\hat{W}$$

[1]The electric susceptibility is defined by $\chi_{i,j} = \frac{1}{\epsilon_0}\frac{dP_i}{dE_j}$, where $\vec{P}$ is the Polarization or electric dipole density and $\vec{E}$ is the electric field. The relative dielectric constant is related to the electric susceptibility by $\epsilon_r = 1 + \chi$. (See ΦSX: Elektrodynamik[32].)

[2]We use the symbol $\hat{W}$ standing for the German word "Wechselwirkung" for Interaction, because the interaction between particles is often treated by perturbation theory.

Thus for $\lambda = 0$ we obtain the unperturbed system and for $\lambda = 1$ the fully perturbed system. For this scaled system, the energy eigenvalues and the wave functions depend on λ.

13.2 General principle of perturbation theory

Let us now consider that we would like to evaluate a quantity F, which can be the energy, the wave function or the density. Since the Hamiltonian depends on the parameter λ, also the quantity F will depend on λ. Thus one can write down a Taylor expansion for $F(\lambda)$ as

$$F(\lambda) = \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{d^n F}{d\lambda^n} \right|_{\lambda=0} \lambda^n$$

By inserting $\lambda = 1$ we obtain the result for the perturbed system.

$$F(1) = \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{d^n F}{d\lambda^n} \right|_{\lambda=0}$$

The first few derivatives are easy to evaluate, but the evaluation of higher derivatives is in general cumbersome. If the perturbation is sufficiently small, higher derivatives also do not contribute much. Therefore, the expression for the quantity F in the perturbed system is approximated by

$$F(1) \approx \sum_{n=0}^{n_{max}} \frac{1}{n!} \left. \frac{d^n F}{d\lambda^n} \right|_{\lambda=0}$$

This is the result for $F(1)$ in perturbation theory of order n_{max}. Typically, n_{max} is chosen equal to one or two, leading to first-order (F^{1st}) and second-order (F^{2nd}) perturbation theory.

$$\begin{aligned} F^{1st}(1) &= F(0) + \left. \frac{dF}{d\lambda} \right|_{\lambda=0} \\ F^{2nd}(1) &= F(0) + \left. \frac{dF}{d\lambda} \right|_{\lambda=0} + \frac{1}{2} \left. \frac{d^2 F}{d\lambda^2} \right|_{\lambda=0} \end{aligned}$$

From the λ-derivative of the Schrödinger equation, we obtain expressions for the λ derivative of the energy-eigenvalues and the wave functions. Higher derivatives can be obtained recursively. Then we use these derivatives to set up a Taylor-series expansion of energies and wave functions in λ. This Taylor-series expansion is truncated, usually already after the linear term, and the truncated series is used as approximation for the values at $\lambda = 1$.

Our derivation, shown in the following, deviates from those in most text books, in that we first determine the exact λ dependent first derivatives and then form higher derivatives from those expressions instead of forming the first and higher derivatives for $\lambda = 0$. The reason for this choice is two-fold. First it is a little simpler to construct the higher derivatives from the first derivative, which we need to evaluate anyway, as compared to determining the second derivative from scratch. More importantly, the λ dependent derivatives can be used to obtain the correct result by integration. This procedure is at times used in numerical evaluations of the perturbed energies and eigenstates, and it is also underlies a number of analytical derivations. The latter, however, are not worked out in this text.

13.3 Time-independent perturbation theory

Here we show the procedure in detail: We consider now the λ-dependent, orthonormal eigenstates $|\psi(\lambda)\rangle$ of the perturbed Hamiltonian $\hat{H}_0 + \lambda\hat{W}$. Since the Schrödinger equation

$$\Big[\underbrace{\hat{H}_0 + \lambda\hat{W}}_{\hat{H}(\lambda)} - E_n(\lambda)\Big]|\psi_n(\lambda)\rangle = 0$$

must be fulfilled for any value of λ, we can also form its derivative with respect to λ.

$$\Big[\hat{H}(\lambda)-E_n(\lambda)\Big]|\psi'_n(\lambda)\rangle+\Big[\hat{W}-E'_n(\lambda)\Big]|\psi_n(\lambda)\rangle=0$$

Here we use $E'_n=\frac{dE_n}{d\lambda}$ and $|\psi'_n\rangle=\frac{d|\psi_n\rangle}{d\lambda}$.

Next we multiply from the left with $\langle\psi_m(\lambda)|$, which corresponds to an eigenstate of the perturbed Hamiltonian.

$$\begin{aligned}0&=\langle\psi_m(\lambda)|\hat{H}(\lambda)-E_n(\lambda)|\psi'_n(\lambda)\rangle+\langle\psi_m(\lambda)|\Big[\hat{W}-E'_n(\lambda)\Big]|\psi_n(\lambda)\rangle\\&=\underbrace{\langle\psi_m(\lambda)|E_m(\lambda)}_{\langle\psi_m(\lambda)|\hat{H}(\lambda)}-E_n(\lambda)|\psi'_n(\lambda)\rangle+\langle\psi_m(\lambda)|\hat{W}-E'_n(\lambda)|\psi_n(\lambda)\rangle\end{aligned}\tag{13.1}$$

For $m=n$ the first term vanishes, so that Eq. 13.1 yields an expression for E'_n:

DERIVATIVE OF THE ENERGY EIGENVALUE WITH RESPECT TO THE STRENGTH OF THE PERTURBATION

$$E'_n(\lambda)\stackrel{\text{def}}{=}\frac{dE_n}{d\lambda}=\langle\psi_n(\lambda)|\hat{W}|\psi_n(\lambda)\rangle\tag{13.2}$$

Thus the changes of the eigenvalues are obtained from the expectation value of the perturbation operator alone.

Let us now consider Eq. 13.1 with $m\neq n$.

$$\begin{aligned}\langle\psi_m(\lambda)|E_m(\lambda)-E_n(\lambda)|\psi'_n(\lambda)\rangle+\langle\psi_m(\lambda)|\hat{W}-E'_n(\lambda)|\psi_n(\lambda)\rangle=0\\ \Rightarrow\qquad\langle\psi_m(\lambda)|\psi'_n(\lambda)\rangle=\frac{\langle\psi_m|\hat{W}|\psi_n(\lambda)\rangle}{E_n(\lambda)-E_m(\lambda)}\end{aligned}\tag{13.3}$$

so that we can express the perturbation of the wave functions as

$$\begin{aligned}|\psi'_n(\lambda)\rangle&=\underbrace{\sum_m|\psi_m(\lambda)\rangle\langle\psi_m(\lambda)|}_{\hat{1}}|\psi'_n(\lambda)\rangle\\&\stackrel{\text{Eq. 13.3}}{=}|\psi_n(\lambda)\rangle\langle\psi_n(\lambda)|\psi'_n(\lambda)\rangle+\sum_{m(m\neq n)}|\psi_m(\lambda)\rangle\frac{\langle\psi_m(\lambda)|\hat{W}|\psi_n(\lambda)\rangle}{E_n(\lambda)-E_m(\lambda)}\end{aligned}$$

This solution is, however, not yet unique, because it still depends on the unknown matrix element $\langle\psi_n|\psi'_n\rangle$. We will show in the following that (1) the real part of $\langle\psi_n|\psi'_n\rangle$ vanishes, and (2) that its imaginary part can be chosen equal to zero.

1. We can use the condition that the wave function is normalized for every value of λ:

$$0=\frac{d}{d\lambda}\underbrace{\langle\psi_n(\lambda)|\psi_n(\lambda)\rangle}_{=1}=\langle\psi'_n(\lambda)|\psi_n(\lambda)\rangle+\langle\psi_n(\lambda)|\psi'_n(\lambda)\rangle=2\mathrm{Re}\{\langle\psi_n(\lambda)|\psi'_n(\lambda)\rangle\}$$

   Thus the real part of $\langle\psi'_n|\psi_n\rangle$ vanishes.

2. What about the imaginary part? It is arbitrary! This is shown as follows: Consider a λ-dependent wave function $|\psi_n(\lambda)\rangle$. For each value of λ the wave function can be multiplied with a phase factor $\mathrm{e}^{i\phi(\lambda)}$. The resulting wave function

$$|\tilde{\psi}_n(\lambda)\rangle=|\psi_n(\lambda)\rangle\mathrm{e}^{i\phi(\lambda)}$$

is an eigenstate for each strength of the perturbation if the same is true for $|\psi_n(\lambda)\rangle$. Now we form the derivative

$$
\begin{aligned}
|\tilde{\psi}'_n(\lambda)\rangle &= \frac{d}{d\lambda}\Big(|\psi_n(\lambda)\rangle e^{i\phi(\lambda)}\Big) = |\psi'_n\rangle e^{i\phi(\lambda)} + |\psi_n\rangle i\phi'(\lambda)e^{i\phi(\lambda)} \\
\Rightarrow \langle\tilde{\psi}_n(\lambda)|\tilde{\psi}'_n(\lambda)\rangle &= e^{-i\phi(\lambda)}\langle\psi_n(\lambda)|\Big(|\psi'_n(\lambda)\rangle + |\psi_n(\lambda)\rangle i\phi'(\lambda)\Big)e^{i\phi(\lambda)} \\
&= \langle\psi_n|\psi'_n\rangle + \underbrace{\langle\psi_n|\psi_n\rangle}_{=1} i\phi'(\lambda) = \langle\psi_n|\psi'_n\rangle + i\phi'(\lambda)
\end{aligned}
$$

Thus, simply by changing the λ-dependent phase factor $\phi(\lambda)$, the imaginary part of $\langle\tilde{\psi}_n|\tilde{\psi}'_n\rangle$ can be changed into an any value. All these solutions are equally valid. Therefore, we choose the most convenient result by setting the imaginary part of $\langle\tilde{\psi}_n|\tilde{\psi}'_n\rangle$ to zero. In the following we drop the tilde on the wave function.

With the argument just provided, the matrix element $\langle\psi_n|\psi'_n\rangle$ vanishes. Thus, with the appropriate choice of the phase factor, we obtain the final result:

DERIVATIVE OF THE WAVE FUNCTION WITH RESPECT TO THE STRENGTH OF THE PERTURBATION

$$
|\psi'_n(\lambda)\rangle \stackrel{\text{def}}{=} \frac{d|\psi_n(\lambda)\rangle}{d\lambda} = \sum_{m(m\neq n))} |\psi_m(\lambda)\rangle \frac{\langle\psi_m(\lambda)|\hat{W}|\psi_n(\lambda)\rangle}{E_n(\lambda) - E_m(\lambda)} \tag{13.4}
$$

Note that Eqs. 13.2 and 13.4 are exact.

13.3.1 Degenerate case

Eq. 13.4 becomes problematic if E_n is degenerate, because, in this case, terms with a divide-by-zero appear in the expression for the wave function.

Without limitation of generality, we assume that the degeneracy is present at $\lambda = 0$. This is also the most relevant case.

The remedy is the following recipe:

1. Determine all the states $|\Psi_n(0)\rangle$ within one multiplet, that is $n \in M$ where M is the set of indices of the states in the same multiplet. A multiplet is the set of degenerate states, that is $E_n(0) = E_m(0)$ for all $n, m \in M$.

2. determine the matrix elements of the perturbation with the states in the multiplet

$$
W_{m,n} = \langle\psi_m(0)|\hat{W}|\psi_n(0)\rangle \qquad \text{with } n, m \in M
$$

and diagonalize it so that

$$
\boldsymbol{W}\vec{c}_j = \vec{c}_j w_j \qquad \text{or} \qquad \sum_m W_{n,m}c_{m,j} = c_{n,j}w_j
$$

The eigenvalues are w_j and the eigenvectors are $\vec{c}_j$.

3. Form a new basisset so that

$$
|\tilde{\psi}_j(0)\rangle = \sum_m |\psi_m(0)\rangle c_{m,j}
$$

4. Now we set up the equation Eq. 13.1 that we obtained before,with the difference that we use the new states with a tilde.

$$\langle\tilde\psi_m|\tilde\psi'_n\rangle(\epsilon_m-\epsilon_n)+\langle\tilde\psi_m|\hat W|\tilde\psi_n\rangle-\langle\tilde\psi_m|\tilde\psi_n\rangle\epsilon'_n=0$$

- Case 1: $\epsilon_m=\epsilon_n$

$$\begin{aligned}\langle\tilde\psi_m|\hat W|\tilde\psi_n\rangle-\langle\tilde\psi_m|\tilde\psi_n\rangle\epsilon'_n&=0\\ \Rightarrow\qquad w_n\delta_{m,n}-\epsilon'_n\delta_{m,n}&=0\\ \Rightarrow\qquad \epsilon'_n&=w_n\end{aligned}$$

- Case 2:

$$|\tilde\psi'_n\rangle=\sum_{m\in M}|\tilde\psi_m\rangle\langle\tilde\psi_m|\tilde\psi'_n\rangle+\sum_{m\notin M}|\tilde\psi_m\rangle\frac{\langle\tilde\psi_m|\hat W|\tilde\psi_n\rangle}{\epsilon_n-\epsilon_m}$$

Next we need to show that the matrix elements $\langle\tilde\psi_m|\tilde\psi'_n\rangle$ can be chosen equal to zero. This is shown as follows: we begin with the λ-derivative of the normalization condition

$$\begin{aligned}\langle\tilde\psi_m|\tilde\psi_n\rangle&=\delta_{m,n}\\ \stackrel{\partial_\lambda}{\Rightarrow}\qquad\langle\tilde\psi'_m|\tilde\psi_n\rangle+\langle\tilde\psi_m|\tilde\psi'_n\rangle&=0\\ \Rightarrow\qquad\langle\tilde\psi'_m|\tilde\psi_n\rangle&=-\langle\tilde\psi'_n|\tilde\psi_m\rangle\end{aligned}$$

This shows that the matrix $\langle\tilde\psi_m|\tilde\psi'_n\rangle$ is anti-hermitian. In the following we proceed analogously to what has been done before by adding a lambda-dependent phase factor. Here, however we need to generalize the argument to several states.
We consider a unitary transformation $\boldsymbol{U}(\lambda)=\mathrm{e}^{\boldsymbol{A}(\lambda)}$ of the states in the multiplet,

$$|\bar{\tilde\psi}_n(\lambda)\rangle=\sum_m|\tilde\psi_m(\lambda)\rangle U_{m,n}(\lambda)=\sum_m|\tilde\psi_m(\lambda)\rangle\left(\mathrm{e}^{\boldsymbol{A}(\lambda)}\right)_{m,n}$$

where $\boldsymbol{A}(\lambda)$ is a antihermitian matrix, that is $\boldsymbol{A}^\dagger=-\boldsymbol{A}$.
It can be shown as follows, that $\boldsymbol{U}:=\mathrm{e}^{\boldsymbol{A}}$ is unitary, if $\boldsymbol{A}$ is antihermitian:

$$\boldsymbol{U}^\dagger\boldsymbol{U}=\left(\mathrm{e}^{\boldsymbol{A}}\right)^\dagger\mathrm{e}^{\boldsymbol{A}}=\mathrm{e}^{\boldsymbol{A}^\dagger}\mathrm{e}^{\boldsymbol{A}}=\mathrm{e}^{-\boldsymbol{A}}\mathrm{e}^{\boldsymbol{A}}=\mathbf{1}$$

We require that the transformation is the identity for $\lambda=0$, that is $\boldsymbol{U}(\lambda=0)=\mathbf{1}$ or $\boldsymbol{A}(\lambda=0)=\mathbf{0}$.
Now we evaluate our expression

$$\begin{aligned}|\bar{\tilde\psi}'_n(\lambda)\rangle&=|\tilde\psi'_m(\lambda)\rangle\left(\mathrm{e}^{\boldsymbol{A}(\lambda)}\right)_{m,n}+|\tilde\psi_m(\lambda)\rangle\left(\boldsymbol{A}\mathrm{e}^{\boldsymbol{A}(\lambda)}\right)_{m,n}\\ \stackrel{\boldsymbol{U}(\lambda=0)=\mathbf{1}}{\Rightarrow}|\bar{\tilde\psi}'_n(0)\rangle&=|\tilde\psi'_m(\lambda)\rangle+|\tilde\psi_m(\lambda)\rangle\left(\boldsymbol{A}'\right)_{m,n}\\ \langle\bar{\tilde\psi}_m(0)|\bar{\tilde\psi}'_n(0)\rangle&=\langle\tilde\psi_m(0)|\tilde\psi'_n(0)\rangle+\sum_k\underbrace{\langle\tilde\psi_m(\lambda)|\tilde\psi_k(\lambda)\rangle}_{\delta_{m,k}}A'_{k,n}\\ &\stackrel{\lambda=0}{=}\langle\tilde\psi_m(\lambda)|\tilde\psi'_n(\lambda)\rangle+A'_{m,n}\end{aligned}$$

Thus one can always find a λ-dependent, unitary transformation so that the states within the multiplet do not contribute to the derivative.

Thus, for the degenerate result we obtain

DERIVATIVES OF ENERGIES AND WAVE FUNCTIONS IN THE DEGENERATE CASE

$$
\begin{aligned}
\epsilon'_n(\lambda) &= \langle\psi_n(\lambda)|\hat{W}|\psi_n(\lambda)\rangle \\
|\psi'_n(\lambda)\rangle &= \sum_{m(\epsilon_m(\lambda)\neq\epsilon_n(\lambda))} |\psi_m(\lambda)\rangle \frac{\langle\psi_m(\lambda)|\hat{W}|\psi_n(\lambda)\rangle}{\epsilon_n(\lambda)-\epsilon_m(\lambda)}
\end{aligned}
$$

where

$$
\langle\psi_n(\lambda)|\hat{W}|\psi_m(\lambda)\rangle = w_n\delta_{m,n}
$$

for all states in the same multiplet.

13.3.2 First order perturbation theory

Now we use the derivatives Eq. 13.2 and Eq. 13.4 for the unperturbed system, and express the perturbed values as Taylor series.

$$
E_n(\lambda) = \sum_j \frac{1}{j!} \left.\frac{d^j E_n}{d\lambda^j}\right|_{\lambda=0} \lambda^j
$$

This Taylor expansion is evaluated for the full perturbation that is for $\lambda = 1$.

$$
\Rightarrow \qquad E_n(\lambda=1) = \sum_j \frac{1}{j!} \left.\frac{d^j E_n}{d\lambda^j}\right|_{\lambda=0}
$$

Then we do the actual perturbations by truncating the expansion after the n-th-order term in λ for the n-th order perturbation theory. In practice only the first and second order perturbation theory is used frequently. For the first order perturbation theory we calculate

$$
E_n^{\text{1st order}} \approx E_n(\lambda=0) + \left.\frac{dE_n}{d\lambda}\right|_{\lambda=0}
$$

and for the second order perturbation theory we include also the next term

$$
E_n^{\text{2nd order}} \approx E_n(\lambda=0) + \left.\frac{dE_n}{d\lambda}\right|_{\lambda=0} + \frac{1}{2}\left.\frac{d^2E_n}{d\lambda^2}\right|_{\lambda=0}
$$

For the wave function we proceed completely analogously.

FIRST ORDER PERTURBATION THEORY

The energy $E_n(1)$ and the energy eigenstates $|\psi_n(1)\rangle$ for a perturbed Hamiltonian $H(1) = \hat{H}_0 + \hat{W}$ have the form

$$
\begin{aligned}
E_n(1) &= E_n(0) + \langle\psi_n(0)|\hat{W}|\psi_n(0)\rangle + O(\hat{W}^2) \\
|\psi_n(1)\rangle &= |\psi_n(0)\rangle + \sum_{m(\epsilon_m\neq\epsilon_n)} |\psi_m(0)\rangle \frac{\langle\psi_m(0)|\hat{W}|\psi_n(0)\rangle}{E_n(0)-E_m(0)} + O(\hat{W}^2)
\end{aligned}
$$

where $E_n(0)$ are the energy eigenvalues and $|\psi_n(0)\rangle$ are the eigenstates of the unperturbed Hamiltonian $\hat{H}_0$
Note, that special precaution is necessary for degenerate states!

Linear response

Once we have an approximation for the wave function in a certain order, we can evaluate the perturbation not only of the energy but of any other observable up to this order. Note, that some terms, which turn up during squaring the wave functions, need to be dropped. We illustrate the principle on the first-order result

$$
\begin{aligned}
\langle\psi_n(\lambda)|\hat{A}|\psi_n(\lambda)\rangle &= \Big(\langle\psi_n(0)| + \lambda\langle\psi'_n(0)|\Big)\hat{A}\Big(|\psi_n(0)\rangle + |\psi'_n(0)\rangle\lambda\Big) + O(\lambda^2) \\
&= \langle\psi_n(0)|\hat{A}|\psi_n(0)\rangle + \lambda\Big(\langle\psi_n(0)|\hat{A}|\psi'_n(0)\rangle + \langle\psi'_n(0)|\hat{A}|\psi_n(0)\rangle\Big) \\
&\quad + \lambda^2\Big(\lambda\langle\psi'_n(0)|\hat{A}|\psi'_n(0)\rangle\Big) + O(\lambda^2) \\
&= \langle\psi_n(0)|\hat{A}|\psi_n(0)\rangle \\
&\quad + \lambda\Big(\langle\psi_n(0)|\hat{A}|\psi'_n(0)\rangle + \langle\psi'_n(0)|\hat{A}|\psi_n(0)\rangle\Big) + O(\lambda^2)
\end{aligned}
$$

Thus we obtain

$$
\langle\psi_n(1)|\hat{A}|\psi_n(1)\rangle \stackrel{\text{1st order}}{\approx} \langle\psi_n(0)|\hat{A}|\psi_n(0)\rangle + \langle\psi_n(0)|\hat{A}|\psi'_n(0)\rangle + \langle\psi'_n(0)|\hat{A}|\psi_n(0)\rangle
$$

LINEAR RESPONSE

Thus the linear response δA_n of an expectation value A_n of an observable $\hat{A}$ in an eigenstate $|\psi_n\rangle$ of the Hamiltonian due to a perturbation $\hat{W}$ is obtained by the following expression.

$$
\delta\hat{A}_n \stackrel{\text{def}}{=} \langle\psi_n(1)|\hat{A}|\psi_n(1)\rangle - \langle\psi_n(0)|\hat{A}|\psi_n(0)\rangle = 2\mathrm{Re}\left(\sum_{m(m\neq n)} \frac{\langle\psi_n|\hat{A}|\psi_m\rangle\langle\psi_m|\hat{W}|\psi_n\rangle}{E_n - E_m}\right)
$$

To give an example, this expression may be useful to determine the dipole of a molecular molecule that is induced by an external electrical field. The proportionality constant is the **polarization** of the molecule.

13.3.3 Second order perturbation theory

The second-order result for the energy eigenvalues is sufficiently simple so that we show it here. For higher order terms even perturbation theory becomes increasingly difficult so that it is rarely done.

The second derivatives of $E_n(\lambda)$ and $|\psi_n(\lambda)\rangle$ are obtained from Eq. 13.2 and Eq. 13.4, which are valid for any value of λ, by deriving them once more with respect to λ. The resulting expression contains first derivatives that are known already. Finally, we insert the result into the Taylor expansion up to second order about $\lambda = 0$ and set $\lambda = 1$

$$
\begin{aligned}
E'_n &= \langle\psi_n|\hat{W}|\psi_n\rangle \\
E''_n &= \langle\psi'_n|\hat{W}|\psi_n\rangle + \langle\psi_n|\hat{W}|\psi'_n\rangle \stackrel{\text{Eq. 13.4}}{=} 2\sum_{m(\neq n)} \frac{|\langle\psi_m|\hat{W}|\psi_n\rangle|^2}{E_n - E_m} \\
E_n &= \bar{E}_n + E'_n + \frac{1}{2}E''_n \\
&= \bar{E}_n + \langle\bar{\psi}_n|\hat{W}|\bar{\psi}_n\rangle + \sum_{m(\neq n)} \frac{|\langle\bar{\psi}_m|\hat{W}|\bar{\psi}_n\rangle|^2}{\bar{E}_n - \bar{E}_m}
\end{aligned}
$$

Note that the sum extends over all states $|\psi_m\rangle$ except $|\psi_n\rangle$. (The index n is fixed.)

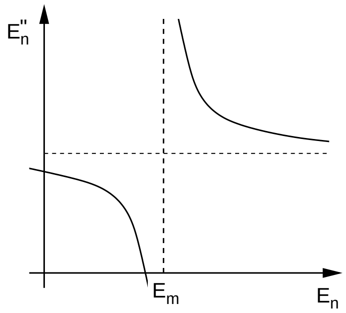


Fig. 13.1: Second order correction E_n'' by a second level E_m. The result of the second order correction is a "level repulsion". It is evident that the correction diverges for $E_n = E_m$ indicating the breakdown of the perturbation series for degenerate energy levels.

It may be interesting to consider the effect of the second order correction of the energy levels. The contribution of nearby energy levels is largest, given that the matrix elements are roughly of similar size.

13.4 Time-dependent perturbation theory

The prototypical application of **time-dependent perturbation theory** is the determination of optical absorption spectra. The light-wave creates an oscillating electric field at the site of the transition, and will induce transitions of electrons from an occupied orbital into an un-occupied orbital. Typically, the electric field is sufficiently small to be treated as a small perturbation. The wave length is usually sufficiently large that we only need to consider a fluctuating, but spatially constant electric field at the point of the transition. For monochromatic light, the perturbation of the Hamiltonian due to the light has the following form

$$\hat{W}(t) = -\hat{\vec{r}}q\vec{E}_0 \sin(\omega t) \tag{13.5}$$

where q is the charge of the excited electron, $\vec{E}_0$ the amplitude of the electric field in the light wave, and $\hbar\omega$ the photon energy.

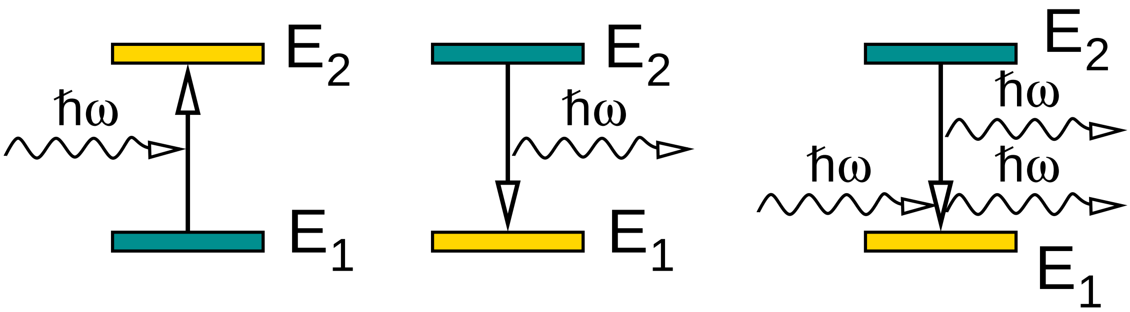


Fig. 13.2: Sketch of time dependent processes. Left **optical absorption**: A photon is absorbed and an electron is lifted to a higher energy level. Right: **Stimulated emission**: If the higher energy level is occupied and the lower level is filled, a photon can trigger the desorption of another one, while the electron drops into the lower energy level. Stimulated emission underlies the workings of a laser. Middle: **Spontaneous emission**: An electron can drop into the lower level while emitting a photon. This process is similar to stimulated emission, but the photon that triggers the emission is a virtual photon, i.e. a vacuum fluctuation of the electromagnetic field.

13.4.1 Interaction picture

In the following, we do not specialize on monochromatic light, but treat a general time-dependent perturbation $\hat{W}(t)$ of the unperturbed and time-independent Hamiltonian $\hat{H}_0$.

$$i\hbar\partial_t|\psi\rangle = \Big(\hat{H}_0 + \hat{W}(t)\Big)|\psi\rangle \tag{13.6}$$

The Schrödinger equation Eq. 13.6 is expressed in the –conventional– Schödinger picture. For perturbation theory, the **interaction picture** is more convenient, because it allows us to "hide away" the wave-function oscillations of the unperturbed system. The interaction picture has been discussed earlier in section 8.5.3 on p. 136. The term *"interaction picture"* is misleading, because we want to study a time-dependent perturbation of a non-interacting system. The choice has a historical reason: The interaction picture has been developed to study the interaction between electrons in perturbation theory. This is also the origin of the subscript W for the quantities in the interaction picture. It is derived from the german word *"Wechselwirkung"* for Interaction.

I indicate the wave functions $|\psi_S\rangle$ of the Schrödinger picture by a subscript S, and those, $|\psi_W\rangle$, of the interaction picture by a subscript W. The wave function in the interaction picture is

$$|\psi_W(t)\rangle \stackrel{\text{def}}{=} \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}|\psi_S(t)\rangle \tag{13.7}$$

Insertion of Eq. 13.7 into Eq. 13.6 yields.

$$\begin{aligned}
i\hbar\partial_t|\psi_W(t)\rangle &\stackrel{\text{Eq. 13.7}}{=} \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}|\psi_S(t)\rangle \\
&= -\hat{H}_0\mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}|\psi_S(t)\rangle + \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t} i\hbar\partial_t|\psi_S(t)\rangle \\
&\stackrel{\text{Eq. 13.6}}{=} -\hat{H}_0\mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}|\psi_S(t)\rangle + \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\Big(\hat{H}_0 + \hat{W}(t)\Big)|\psi_S(t)\rangle \\
&= \underbrace{\mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\overbrace{\hat{W}(t)}^{\hat{W}_S(t)}\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t}}_{\hat{W}_W(t)}|\psi_W(t)\rangle
\end{aligned} \tag{13.8}$$

Thus the time-dependent Schrödinger equation in the interaction picture has the simple form

TIME-DEPENDENT SCHRÖDINGER EQUATION IN THE INTERACTION PICTURE

$$i\hbar\partial_t\big|\psi_W(t)\big\rangle = \hat{W}_W(t)\big|\psi_W(t)\big\rangle \tag{13.9}$$

with

$$|\psi_W(t)\rangle \stackrel{\text{def}}{=} \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}|\psi_S(t)\rangle \qquad \text{and} \qquad \hat{W}_W(t) \stackrel{\text{def}}{=} \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\hat{W}_S(t)\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t} \tag{13.10}$$

13.4.2 Perturbation theory

First, we turn the Schrödinger equation Eq. 13.9 in the interaction picture into an integral equation. Without restriction of generality, we assume that the interaction vanishes before the time t_0, i.e. $\hat{W}(t) = \hat{0}$ for $t \le t_0$.

$$\big|\psi_W(t)\big\rangle = \big|\psi_W(t_0)\big\rangle - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \lambda\hat{W}_W(t')\big|\psi_W(t')\big\rangle \tag{13.11}$$

Unfortunately, the solution occurs again on both sides of the equation. Thus, it is an equation that needs to be solved iteratively. Perturbation theory provides such an iterative scheme.

We introduce a scale parameter λ for the interaction, which produces an equation for the λ-dependent wave functions.

$$\big|\psi_W(\lambda,t)\big\rangle = \big|\psi_W(t_0)\big\rangle - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \lambda\hat{W}_W(t')\big|\psi_W(\lambda,t')\big\rangle \tag{13.12}$$

Next, we express the λ-dependent wave function by a Taylor expansion

$$|\psi_W(\lambda,t)\rangle = \sum_{n=0}^{\infty}\frac{1}{n!}|\psi_W^{(n)}\rangle\lambda^n \tag{13.13}$$

and insert it into the equation above. We sort the terms in orders of λ.

$$\begin{aligned}
\frac{1}{n!}\big|\psi_W^{(n)}(t)\big\rangle &= \frac{1}{n!}\big|\psi_W^{(n)}(t_0)\big\rangle\delta_{n,0} - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{W}_W(t')\frac{1}{(n-1)!}\big|\psi_W^{(n-1)}(t')\big\rangle \\
\Leftrightarrow \qquad \big|\psi_W^{(n)}(t)\big\rangle &= \big|\psi_W^{(n)}(t_0)\big\rangle\delta_{n,0} - n\frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{W}_W(t')\big|\psi_W^{(n-1)}(t')\big\rangle
\end{aligned} \tag{13.14}$$

We introduced the Kronecker delta, because the wave function at t_0 is independent of the perturbation due to the requirement, $\hat{W}(t) = \hat{0}$ for $t \le t_0$, specified above.

Thus we obtain[3] the first few λ-derivatives as

$$\begin{aligned}
\big|\psi_W^{(0)}(t)\big\rangle &= \big|\psi_W^{(0)}(t_0)\big\rangle \\
\big|\psi_W^{(1)}(t)\big\rangle &= -\frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{W}_W(t')\big|\psi_W^{(0)}(t')\big\rangle = -\frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{W}_W(t')\big|\psi_W^{(0)}(t_0)\big\rangle \\
\big|\psi_W^{(2)}(t)\big\rangle &= -2\frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{W}_W(t')\big|\psi_W^{(1)}(t')\big\rangle = 2\left(-\frac{i}{\hbar}\right)^2\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\ \hat{W}_W(t')\hat{W}_W(t'')\big|\psi_W^{(0)}(t_0)\big\rangle
\end{aligned} \tag{13.15}$$

These expressions can be combined in the second-order expression

$$\begin{aligned}
|\psi_W(\lambda,t)\rangle &\stackrel{\text{Eq. 13.13}}{=} \big|\psi_W^{(0)}(t_0)\big\rangle + \underbrace{\left(-\frac{i\lambda}{\hbar}\right)\int_{t_0}^{t} dt'\ \hat{W}_W(t')\big|\psi_W^{(0)}(t_0)\big\rangle}_{|\psi_W^{(1)}\rangle\lambda} \\
&+ \underbrace{\left(-\frac{i\lambda}{\hbar}\right)^2\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\ \hat{W}_W(t')\hat{W}_W(t'')\big|\psi_W^{(0)}(t_0)\big\rangle}_{\frac{1}{2}|\psi_W^{(2)}\rangle\lambda^2} + O(\lambda^3)
\end{aligned} \tag{13.16}$$

[3] If there is an equation of the form

$$F(t,\lambda) = G(t,\lambda)$$

which is fulfilled for all values of λ, we can form the derivatives.

$$\frac{\partial^n F(t,\lambda)}{\partial\lambda^n} = \frac{\partial^n G(t,\lambda)}{\partial\lambda^n}$$

Since all the derivatives of left and right-hand side are equal, also every term of the Taylor expansion

$$\frac{1}{n!}\frac{\partial^n F(t,\lambda)}{\partial\lambda^n} = \frac{1}{n!}\frac{\partial^n G(t,\lambda)}{\partial\lambda^n}$$

and the truncated Taylor expansions must be equal.

$$\sum_{n=0}^{k}\frac{1}{n!}\frac{\partial^n F(t,\lambda)}{\partial\lambda^n}\lambda^k = \sum_{n=0}^{k}\frac{1}{n!}\frac{\partial^n G(t,\lambda)}{\partial\lambda^n}\lambda^k$$

This is a fundamental step done in perturbation theory.

In the last step, we set $\lambda = 1$ in the truncated expansion, which yields the perturbed wave function to second order. In addition, we transform the result back into the Schrödinger picture using Eq. 13.10.

WAVE FUNCTION IN SECOND ORDER PERTURBATION THEORY

$$
\begin{aligned}
|\psi_W(t)\rangle &= \big|\psi_W^{(0)}(t_0)\big\rangle + \left(-\frac{i}{\hbar}\right)\int_{t_0}^{t} dt'\ \hat{W}_W(t')\big|\psi_W^{(0)}(t_0)\big\rangle \\
&\quad + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\ \hat{W}_W(t')\hat{W}_W(t'')\big|\psi_W^{(0)}(t_0)\big\rangle + O(\hat{W}^3)
\end{aligned}
\tag{13.17}
$$

We went to second order, because we often need the perturbation of an expectation value. For these expressions the first order often vanishes and the second order expression requires both the first and the second order in the wave function.

In the Schrödinger picture Eq. 13.17 has the form

$$
\begin{aligned}
|\psi_S(t)\rangle &= \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0(t-t_0)}\big|\psi_S(t_0)\big\rangle + \left(-\frac{i}{\hbar}\right)\int_{t_0}^{t} dt'\ \mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0(t-t')}\hat{W}_S(t')\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0(t'-t_0)}\big|\psi_S(t_0)\big\rangle \\
&\quad + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\ \mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0(t-t')}\hat{W}_S(t')\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0(t'-t'')}\hat{W}_S(t'')\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0(t''-t_0)}\big|\psi_S(t_0)\big\rangle \\
&\quad + O\left(\hat{W}^3\right)
\end{aligned}
\tag{13.18}
$$

The picture, which emerges in Eq. 13.17 is that of a unperturbed particle, which evolves in time, but which scatters once (in the first order) or twice (in the second order) at the perturbation. By adding all pathways with zero, one, two etc. scattering events, the final wave function is obtained.

Dyson's time ordering operator (optional)

Let me now take a little detour, which prepares the ground for techniques used in many-particle physics. Namely, we will seek an expression to write down the solution to arbitrary order in the interaction.

Let me return to the second derivative of the wave function in λ. The integration bounds are set so that the time arguments of the product are always increasing from right to left. In a first step, I change the integration bounds to the full interval $[t_0, t]$ and select the correct terms with the Heaviside function[4] $\theta(t)$

$$
\begin{aligned}
|\psi_W^{(2)}(t)\rangle &= 2\left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\ \hat{W}_W(t')\hat{W}_W(t'')\big|\psi_W^{(0)}(t_0)\big\rangle \\
&= \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \left[2\theta(t'-t)\hat{W}_W(t')\hat{W}_W(t'')\right]\big|\psi_W^{(0)}(t_0)\big\rangle \\
&= \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \Big[\underbrace{\theta(t'-t)\hat{W}_W(t')\hat{W}_W(t'') + \theta(t-t')\hat{W}_W(t'')\hat{W}_W(t')}_{=:\mathcal{T}_D\{\hat{W}_W(t')\hat{W}_W(t'')\}}\Big]\big|\psi_W^{(0)}(t_0)\big\rangle \\
&= \mathcal{T}_D\left\{\left(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{W}_W(t')\right)^2\right\}\big|\psi_W^{(0)}(t_0)\big\rangle
\end{aligned}
\tag{13.19}
$$

In the last step, I introduced **Dyson's time-ordering operator** $\mathcal{T}_D$.

[4]The Heaviside function is a step function which vanishes for negative arguments and which is unity for positive arguments.

DYSON'S TIME-ORDERING OPERATOR

Dyson's time ordering operator[40] arranges the time-dependent operators in a product so that the time is ascending from the right to left.

$$\mathcal{T}_D\Big\{\hat{A}(t)\hat{B}(t')\Big\} = \begin{cases} \hat{A}(t)\hat{B}(t') & \text{for } t > t' \\ \hat{B}(t)\hat{A}(t') & \text{for } t < t' \end{cases} \tag{13.20}$$

Dyson's time ordering operator is not an operator in the conventional sense. It cannot be applied to a wave function. Rather it is simply a rule on how to arrange time dependent operators in a product.

The reader may confirm by himself that the higher derivatives can be constructed analogously, so that

$$|\psi_W^{(n)}(t)\rangle = \mathcal{T}_D\Big\{\Big(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\,\hat{W}_W(t')\Big)^n\Big\}|\psi_W^{(0)}(t_0)\rangle \tag{13.21}$$

This allows to include all orders in the expression for the perturbed wave function

$$\begin{aligned} |\psi_W(\lambda,t)\rangle &= \sum_{n=0}^{\infty}\frac{1}{n!}\mathcal{T}_D\Big\{\Big(-\frac{i\lambda}{\hbar}\int_{t_0}^{t} dt'\,\hat{W}_W(t')\Big)^n\Big\}|\psi_W^{(0)}(t_0)\rangle \\ &= \mathcal{T}_D\Big\{\exp\Big(-\frac{i\lambda}{\hbar}\int_{t_0}^{t} dt'\,\hat{W}_W(t')\Big)\Big\}|\psi_W^{(0)}(t_0)\rangle \end{aligned} \tag{13.22}$$

In the last step, we used the Taylor expansion of the exponential $\mathrm{e}^x = \sum_{n=0}^{\infty}\frac{1}{n!}x^n$.

Dyson's time ordering operator allowed us to write down a closed expression[5] to all orders in perturbation theory.

13.4.3 Transition probabilities

Here, we use perturbation theory to estimate the transition probability between two states by an optical excitation. The basic problem is the following:

1. A system is initially prepared in an eigenstate ϕ_n of an unperturbed hamiltonian

$$|\Psi_S(t_0)\rangle = |\phi_n\rangle \tag{13.23}$$

2. The system then experiences a time-dependent perturbation, which has the form of a pulse with a leading frequency ω and a finite duration

$$\hat{W}_S(t) = \hat{W}_0\mathrm{e}^{-i\omega t}\chi(t)\ , \tag{13.24}$$

where $\chi(t)$ is the **pulse shape**.

3. After the perturbation, i.e. at t_f, we measure an observable $\hat{A} = \sum_m |\phi_m\rangle a_m\langle\phi_m|$, which is a constant of motion of the unperturbed system. In order to predict its value we need to evaluate the probabilities of finding the system in another eigenstate $|\phi_m\rangle$ of the unperturbed Hamiltonian.

$$P_{m\leftarrow n} = \langle\psi_S(t_f)|\phi_m\rangle\langle\phi_m|\psi_S(t_f)\rangle \tag{13.25}$$

[5]We assume that $\lambda = 1$ lies within the convergence radius of the Taylor expansion.

Let us evaluate the expectation value of an observable $\hat{A}$ at time t.

$$
\begin{aligned}
A(t) &\stackrel{\text{def}}{=} \langle\psi_S(t)|\hat{A}_S|\psi_S(t)\rangle = \langle\psi_W(t)|\hat{A}_W(t)|\psi_W(t)\rangle \\
&= \langle\psi_W^{(0)}(t)|\hat{A}_W(t)|\psi_W^{(0)}(t)\rangle + \underbrace{\langle\psi_W^{(1)}(t)|\hat{A}_W(t)|\psi_W^{(0)}(t)\rangle + \langle\psi_W^{(0)}(t)\hat{A}_W(t)\psi_W^{(1)}(t)\rangle}_{\text{1st order}} \\
&\quad + \underbrace{\langle\psi_W^{(2)}(t)|\hat{A}_W(t)|\psi_W^{(0)}(t)\rangle + \langle\psi_W^{(1)}(t)|\hat{A}_W(t)|\psi_W^{(1)}(t)\rangle + \langle\psi_W^{(0)}(t)|\hat{A}_W(t)|\psi_W^{(2)}(t)\rangle}_{\text{2nd order}} + O(\hat{W}^3) \\
&= \langle\psi_W^{(0)}|\hat{A}_W(t)|\psi_W^{(0)}\rangle \\
&\quad + \left(\frac{i}{\hbar}\right)\int_{t_0}^{t} dt' \left(\langle\psi_W^{(0)}|\hat{W}_W(t')\hat{A}_W(t)|\psi_W^{(0)}\rangle - \langle\psi_W^{(0)}|\hat{A}_W(t)\hat{W}_W(t')|\psi_W^{(0)}\rangle\right) \\
&\quad + \left(\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \Big(\langle\psi_W^{(0)}|\theta(t''-t')\hat{W}_W(t')\hat{W}_W(t'')\hat{A}_W(t)|\psi_W^{(0)}\rangle \\
&\quad - \langle\psi_W^{(0)}|\hat{W}_W(t')\hat{A}_W(t)\hat{W}_W(t'')|\psi_W^{(0)}\rangle + \langle\psi_W^{(0)}|\hat{A}_W(t)\hat{W}_W(t')\hat{W}_W(t'')\theta(t'-t'')|\psi_W^{(0)}\rangle\Big) \\
&\quad + O(\hat{W}^3)
\end{aligned}
\tag{13.26}
$$

I dropped the time argument from $|\psi_W^{(0)}\rangle$, because it is time independent.

Let me now introduce the eigenstates $|\phi_n\rangle$ of the unperturbed Hamiltonian $\hat{H}_0$, which have energies E_n.

$$
\hat{H}_0|\phi_n\rangle = |\phi_n\rangle E_n \qquad \text{with} \qquad \langle\phi_m|\phi_n\rangle = \delta_{m,n} \tag{13.27}
$$

It makes sense to measure an observable A that (1) is time independent and that (2) furthermore commutates with the unperturbed Hamiltonian $\hat{H}_0$, i.e.

$$
[\hat{A}_S, \hat{H}_0]_- = 0 \tag{13.28}
$$

so that the measurement does not depend on time, once the perturbation is switched off. We choose this limitation, because it removes a lot of complexity. Then, we can also use

$$
\hat{A}_W(t) = \mathrm{e}^{\frac{i}{\hbar}H_0 t}\hat{A}_S\mathrm{e}^{-\frac{i}{\hbar}H_0 t} = \hat{A}_S = \sum_n |\phi_n\rangle a_n\langle\phi_n| \tag{13.29}
$$

together with the identity operator $\hat{1} = \sum_n |\phi_n\rangle\langle\phi_n|$.

$$
\begin{aligned}
A(t) =& \sum_n \langle\psi_W^{(0)}|\phi_n\rangle a_n \langle\phi_n|\psi_W^{(0)}\rangle \\
&+ \left(\frac{i}{\hbar}\right)\sum_{m,n}\int_{t_0}^{t} dt' \Big(\langle\psi_W^{(0)}|\phi_m\rangle\langle\phi_m|\hat{W}_W(t')|\phi_n\rangle a_n \langle\phi_n|\psi_W^{(0)}\rangle \\
&\qquad\qquad - \langle\psi_W^{(0)}|\phi_m\rangle a_m \langle\phi_m|\hat{W}_W(t')|\phi_n\rangle\langle\phi_n|\psi_W^{(0)}\rangle \Big) \\
&+ \left(\frac{i}{\hbar}\right)^2 \sum_{m,n,f}\int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \Big(\theta(t''-t')\langle\psi_W^{(0)}|\phi_m\rangle\langle\phi_m|\hat{W}_W(t')|\phi_f\rangle\langle\phi_f|\hat{W}_W(t'')|\phi_n\rangle a_n \langle\phi_n|\psi_W^{(0)}\rangle \\
&\qquad - \underbrace{\Big(\theta(t''-t')+\theta(t'-t'')\Big)}_{=1} \langle\psi_W^{(0)}|\phi_m\rangle\langle\phi_m|\hat{W}_W(t')|\phi_f\rangle a_f \langle\phi_f|\hat{W}_W(t'')|\phi_n\rangle\langle\phi_n|\psi_W^{(0)}\rangle \\
&\qquad + \theta(t'-t'')\langle\psi_W^{(0)}|\phi_m\rangle a_m \langle\phi_m|\hat{W}_W(t')|\phi_f\rangle\langle\phi_f|\hat{W}_W(t'')|\phi_n\rangle\langle\phi_n|\psi_W^{(0)}\rangle \Big) + O(\hat{W}^3) \\
=& \sum_{m,n} \underbrace{\langle\phi_n|\psi_W^{(0)}\rangle\langle\psi_W^{(0)}|\phi_m\rangle}_{\rho_{m,n} \quad \text{density matrix}} \Big\{ a_n\delta_{m,n} + \left(\frac{i}{\hbar}\right)(a_n-a_m)\int_{t_0}^{t} dt' \, \langle\phi_m|\hat{W}_W(t')|\phi_n\rangle \\
&\qquad + \left(\frac{i}{\hbar}\right)^2 \sum_f \int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \Big(\theta(t''-t')(a_n-a_f)+\theta(t'-t'')(a_m-a_f)\Big) \\
&\qquad \times \langle\phi_m|\hat{W}_W(t')|\phi_p\rangle\langle\phi_p|\hat{W}_W(t'')|\phi_n\rangle \Big\} + O(\hat{W}^3)
\end{aligned}
\tag{13.30}
$$

We see that the expressions simplify considerably when the wave function is initially in a pure state, so that only terms with $m = n$ contribute. To study the transitions from an eigenstate of the unperturbed system was our original intention.

Let me therefore choose $|\phi_i\rangle$ as initial state. With

$$
|\psi_W^{(0)}\rangle = |\phi_i\rangle \qquad \rightsquigarrow \qquad \rho_{m,n} = \delta_{m,i}\delta_{n,i}
\tag{13.31}
$$

we obtain

$$
\begin{aligned}
A(t) &= a_i + \left(\frac{i}{\hbar}\right)^2 \sum_f (a_i - a_f) \underbrace{\int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \, \langle\phi_i|\hat{W}_W(t')|\phi_f\rangle\langle\phi_f|\hat{W}_W(t'')|\phi_i\rangle}_{\left|\int_{t_0}^{t} dt' \langle\phi_i|\hat{W}_W(t')|\phi_f\rangle\right|^2} + O(\hat{W}^3) \\
&= a_i - \sum_f (a_i - a_f)\left|\frac{1}{\hbar}\int_{t_0}^{t} dt' \langle\phi_i|\hat{W}_W(t')|\phi_f\rangle\right|^2 + O(\hat{W}^3) \\
&= a_i\Big(1 - \sum_p \underbrace{\left|\frac{1}{\hbar}\int_{t_0}^{t} dt' \langle\phi_i|\hat{W}_W(t')|\phi_f\rangle\right|^2}_{\Gamma_{i\leftarrow f}}\Big) + \sum_f a_f \underbrace{\left|\frac{1}{\hbar}\int_{t_0}^{t} dt' \langle\phi_i|\hat{W}_W(t')|\phi_f\rangle\right|^2}_{\Gamma_{f\leftarrow i}} + O(\hat{W}^3)
\end{aligned}
\tag{13.32}
$$

This result reminds us of transitions between states which occur with conditional probabilities $\Gamma_{f\leftarrow i}(t, t_0)$, which we call transition probabilities

$$
P_f(t) = P_f(t_0) + \sum_i \Big(\Gamma_{f\leftarrow i}(t,t_0)P_i(t_0) - \Gamma_{i\leftarrow f}(t,t_0)P_f(t_0)\Big) ,
\tag{13.33}
$$

where the **transition probabilities** are

$$\Gamma_{f\leftarrow i}(t,t_0)=\left|\frac{1}{\hbar}\int_{t_0}^{t}dt'\;\underbrace{\langle\phi_f|\hat{W}_S(t')|\phi_i\rangle\mathrm{e}^{\frac{i}{\hbar}(E_f-E_i)t'}}_{\hat{W}_W(t)}\right|^2+O(\hat{W}_S^3)\;. \tag{13.34}$$

so that $A(t)=\sum_m P_m(t)a_m$.

There are subtleties in Eq. 13.33 that may easily be overlooked: They are related to the fact that

$$\Gamma_{f\leftarrow i}(t,t_0)=\sum_q\Gamma_{f\leftarrow q}(t,t')\Gamma_{q\leftarrow i}(t',t_0)\qquad\textbf{is violated}. \tag{13.35}$$

The identity must hold for a true conditional probability. Because the identity Eq. 13.35 is violated, one cannot construct a differential equation by dividing the time interval into many smaller pieces. Hence, Eq. 13.33 is not equivalent to a **kinetic equation** or a **rate equation**.[6]

The identity Eq. 13.35 is violated in our derivation, because we always start from an eigenstate of the unperturbed system. This implies that we discard the phase information of the initial wave function $|\psi(t_0)\rangle=\sum_m|\phi_m\rangle c_m$. The relative phases of different coefficients c_m are ignored. This approximation is equivalent to averaging over all possible relative phases, so that $c_m^*c_n\to\underbrace{P_n}_{|c_n|^2}\delta_{m,n}$, where P_n is a real-valued probability.[7]

Instead of propagating an initial wave function, we propagate the eigenstates $|\phi_m\rangle$ of the unperturbed Hamiltonian individually and average the results with the probabilities $P_m=|c_m|^2$.

Dividing the time interval in small pieces and using Eq. 13.34 as transition probabilities underlies the Boltzmann equation.

Outlook on the Boltzmann equations

The result that we can describe the dynamics by transition probabilities is completely central question in physics. It introduces a time arrow in physics. While the microscopic equations of motion are time-reversible, most complex processes in nature appear to evolve only in one direction.

Let me dwell a little on this subject. This description uses concepts that have not been described and it does not attempt to be rigorous.

In the **Boltzmann equation**, the transport of particles is described in terms of classical particles that scatter at obstacles. The scattering event is described in terms of the conditional probabilities calculated above. The Boltzmann equation completely discards the information about the phase of the wave function. This implies that a lot of quantum-mechanical information is lost. The derivation of the Boltzmann equation is thus right at the border between a classical and a quantum description of nature.

The approximation underlying the Boltzmann equatuion can be described (not shown here) as a randomization of the relative phases of the wave function coefficients before each scattering event.

There is a puzzle in statistical physics, namely that the Schrödinger equation is deterministic, and therefore cannot describe that **entropy** grows. This seems to imply that a quantum system cannot come into **thermal equilibrium**, which is the basic assumption of thermodynamics. In the Boltzmann

[6]Side remark: In order to convert Eq. 13.33 into a kinetic equation, we write it in vector matrix form

$$\vec{P}(t)=\boldsymbol{M}(t,t_0)\vec{P}(t_0)\qquad\text{with } M_{f,i}(t,t_0)=\delta_{f,i}\Big(1-\sum_n\Gamma_{n\leftarrow f}(t,t_0)\Big)+\Gamma_{f\leftarrow i}(t,t_0)$$

$$\Rightarrow\qquad\partial_t\vec{P}(t)=\Big(\partial_t\boldsymbol{M}(t,t_0)\Big)\underbrace{\boldsymbol{M}^{-1}\vec{P}(t)}_{\vec{P}(t_0)}=\Big(\partial_t\ln[\boldsymbol{M}(t,t_0)]\Big)\vec{P}(t) \tag{13.36}$$

If the matrix $\partial_t\ln(\boldsymbol{M})$ becomes constant for $t>t_c$, we can use it to extract the transition rates.

[7]A probablility is real-valued and obeys $P_n\geq 0$, $\sum_n P_n=1$.

equation, on the other hand, the system relaxes in an equilibrium state and the entropy grows in this process.

What allows the entropy to change in the Boltzmann equation? The entropy is related to the amount of uncertainty about a system. As we drop the phase information in between scattering events, we loose information and thus the entropy grows. This shows, that the entropy growth in the Boltzmann equation is an result of its inherent approximations.

The puzzle, that we observe that systems approach thermal equilibrium, while quantum mechanics does not allow for it, is still open. In order to understand entropy growth, we need to consider the environment of the system, i.e. we need to consider **open systems**: When we divide the universe into the system under study and the rest, the environment, we notice that the entropy of the universe is smaller than the sum of the entropies of system and environment.

$$S(A+B) < S(A) + S(B) \tag{13.37}$$

The reason lies in the definition of the entropy: The information on the relative quantum mechanical phases in systems A and B is considered in $S(A+B)$, but not in either $S(A)$ nor in $S(B)$. Thus the entropy of a system grows, because the phase information of the states in the system and the environment gets lost.

Perturbation theory and sum rules: Let me alert you to one aspect of the derivation of the transition probability Eq. 13.34 that is easily overlooked:

1. There is an important sum-rule, namely the sum over all probabilities must be equal to one. This requires the sum over conditional probabilities to add up to one.
2. This sum rule is satisfied only, when the perturbation expansion is carries through consistently up to a given order in the matrix element. The argument is that the sum rule is obeyed for any value of the the perturbation strength λ. Hence, any derivative of the sum rule with λ vanishes. Thus, every expansion to a given order in λ obeys the sum rule.
3. If we had only calculated the perturbation of the amplitude to first order and squared the result, we would not have included all second order terms in the matrix element. Thus the sum-rule would be violated. In particular, we would have gotten all transitions, but we would have missed the reduction of the probability to remain in the initial state.

13.4.4 Transition probability for a pulse

So far, we investigated the time evolution during the pulse. Let me now consider the transition probability for a pulse with finite duration.

The pulse shall have a dominant frequency ω_0 and a finite duration, expressed by a **pulse shape** $\chi(t)$

$$\hat{W}(t) = \hat{W}_0 \mathrm{e}^{-i\omega_0 t}\chi(t) \tag{13.38}$$

The Fourier transform[8] of the pulse-shape is, according to Eqs. Q.4 andQ.3,

$$\chi(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, \tilde{\chi}(\omega)\mathrm{e}^{-i\omega t} \qquad \text{and} \qquad \tilde{\chi}(\omega) = \int_{-\infty}^{\infty} dt\, \chi(t)\mathrm{e}^{+i\omega t} \tag{13.39}$$

Because the pulse shape function $\chi(t)$ is defined to be unit free, its Fourier transform $\tilde{\chi}(\omega)$ has the unit *"time"*.

[8]We change the sign for ω in the definition of the Fourier transform. This is because we usually write a wave as $\mathrm{e}^{i(kx-\omega t)}$.

Insertion of the pulse Eq. 13.38 into Eq. 13.34 for $t_0 = -\infty$ and $t_f = +\infty$

$$\Gamma_{f\leftarrow i} \stackrel{\text{Eq. 13.34}}{=} \left| \frac{1}{\hbar} \int_{-\infty}^{\infty} dt \; \underbrace{\langle \phi_f | \hat{W}_S(t) | \phi_i \rangle e^{\frac{i}{\hbar}(E_f - E_i)t}}_{\hat{W}_W(t)} \right|^2 + O(\hat{W}_S^3) \; . \tag{13.40}$$

shows that the transition probability is determined by the Fourier transform of the pulse shape function.

TRANSITION PROBABILITY

The transition probability from state n to state m induced by perturbation $\hat{W}(t) = \hat{W}_0 e^{-i\omega_0 t}\chi(t)$ with the pulse-shape function $\chi(t)$ is

$$\Gamma_{f\leftarrow i} = \left| \frac{1}{\hbar} \langle \phi_f | \hat{W}_0 | \phi_i \rangle \tilde{\chi} \left(\frac{1}{\hbar}(E_f - E_i - \hbar\omega_0) \right) \right|^2 + O(\hat{W}_0^3) \tag{13.41}$$

where $\tilde{\chi}(\omega) = \int dt \; \chi(t) e^{-i\omega t}$ is the Fourier transform of the pulse-shape function as defined in Eq. 13.39. Note, that the pulse shape function $\chi(t)$ is unit-free. The Fourier transform $\tilde{\chi}(\omega)$ has the unit of *"time"*. The transition probability is thus unit-less. Note also, that the definition of the Fourier transform may differ by a factor, which is absorbed in the prefactor.

A general property of a Fourier transform is that a short pulse automatically leads to a poor frequency resolution. Therefore, short pulses excite over a wide frequency range. A monochromatic pulse is necessarily of infinite duration.

13.4.5 Fermi's golden rule

Let us now investigate excitations with a long, and nearly feature-less, pulse shape. This will lead us to **Fermi's golden rule**, in which the light is nearly mono-chromatic and where shape of the pulse does no more show up.

Let me form a sequence of pulses, Eq. 13.38,

$$\hat{W}_\lambda(t) = \hat{W}_0 e^{-i\omega_0 t} \chi_\lambda(t) \tag{13.42}$$

with

$$\chi_\lambda(t) = \sqrt{\lambda}\chi_1(\lambda t) \; . \tag{13.43}$$

The pulse duration grows to infinity while λ approaches zero.

$$\tau(\lambda) = \sqrt{ \left(\frac{\int dt \, |\chi_\lambda(t)|^2 t^2}{\int dt \, |\chi_\lambda(t)|^2} \right) - \left(\frac{\int dt \, |\chi_\lambda(t)|^2 t}{\int dt \, |\chi_\lambda(t)|^2} \right)^2 } = \frac{1}{\lambda}\tau(1) \tag{13.44}$$

The Fourier transform of the scaled pulse shape can be obtained from that of the unscaled one.

$$\tilde{\chi}_\lambda(\omega) \stackrel{\text{Eq. 13.39}}{=} \frac{1}{\sqrt{\lambda}} \tilde{\chi}_1 \left(\frac{1}{\lambda}\omega \right) \tag{13.45}$$

The norm of a function in the time and in the frequency domain are related by[9] Eq. Q.7 determined in section Q.5.1 on p. 345.

$$\int_{-\infty}^{\infty} dt \; \chi_\lambda^*(t)\chi_\lambda(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \; \tilde{\chi}_\lambda^*(\omega)\tilde{\chi}_\lambda(\omega) \tag{13.46}$$

[9]Caution: The Fourier transform can be defined with different factors. This affects the norm of the Fourier transform. We use here the definitions Eqs. Q.4 andQ.3

A sequence of normalized functions, which approaches zero everywhere, except for a point, converges to Dirac's delta-function. Thus, the square of the pulse shape is related to a delta function in the limit of a monochromatic light pulse.

$$|\tilde{\chi}_\lambda(\omega)|^2 \stackrel{\lambda\to 0}{\longrightarrow} 2\pi\delta(\omega-\omega_0)\int dt\,|\chi_1(t)|^2 \tag{13.47}$$

When we insert this result Eq. 13.47 into Eq. 13.41, we obtain the special case of monochromatic light.

$$\begin{aligned}\Gamma_{f\leftarrow i} &\stackrel{\langle(\omega-\omega_0)^2\rangle\to 0}{\longrightarrow} \frac{2\pi}{\hbar^2}|\langle\phi_f|\hat{W}_0|\phi_i\rangle|^2\delta\left(\omega_0-\frac{1}{\hbar}(E_f-E_i)\right)\int dt\,|\chi_1(t)|^2\\ &= \frac{2\pi}{\hbar}|\langle\phi_f|\hat{W}_0|\phi_i\rangle|^2\delta\left(E_f-E_i-\hbar\omega_0\right)\int dt\,|\chi_1(t)|^2\end{aligned} \tag{13.48}$$

The **transition probability** due to a monochromatic pulse with shape $\hat{W}(t)=\hat{W}_0\mathrm{e}^{-i\omega_0 t}\chi(t)$ with a smooth pulse shape function $\chi(t)$ is

$$\Gamma_{f\leftarrow i} \stackrel{\text{Eq. 13.48}}{=} \frac{2\pi}{\hbar}\left|\langle\phi_f|\hat{W}_0|\phi_i\rangle\right|^2\left(\int dt\,|\chi_1(t)|^2\right)\delta\left(E_f-E_i-\hbar\omega_0\right) \tag{13.49}$$

Observe that the pulse shape only enters in the form of the norm. The shape itself became irrelevant.

This equation says that an excitation with monochromatic light is only possible if the incident light beam has exactly that energy $\hbar\omega_0$ that is necessary to lift the system into a higher-, or lower-energetic state. This requirement is a variant of the energy-conservation law: Even though our light wave is a classical field, the excitations behave as if descrete particles with energy $\hbar\omega_0$ would be absorbed. These particles are **photons**. The relevance of this will become clear, in quantum-field theory, where the amplitude of the light field and the wave function are quantized as well.

Many pulses

This result Eq. 13.51 can be converted into one that gives the transition probability per unit time, rather than the probability for a single pulse. We consider the effect of many light pulses, each described by its own envelope function $\chi_k(t)$ so that

$$\frac{1}{T}\int_0^T dt\,\sum_k|\chi_k(t)|^2=1 \tag{13.50}$$

Then we add up the probabilities from all pulses and obtain the corresponding transition rate $\Lambda_{f\leftarrow i}$. This is **Fermi's golden rule**

FERMI'S GOLDEN RULE

The **transition rate** (transition probability per unit time) due to approximately monochromatic pulses with shape $\hat{W}(t)=\hat{W}_0\mathrm{e}^{-i\omega_0 t}\chi_k(t)$, having a smooth pulse shape $\chi_k(t)$ and a total intensity $\frac{1}{T}\int_0^T dt\,\sum_k|\chi_k(t)|^2=1$ is

$$\Gamma_{f\to i}=\frac{1}{T}\Gamma_{f\to i}\stackrel{\text{Eq. 13.51}}{=}\frac{2\pi}{\hbar}|\langle\phi_f|\hat{W}_0|\phi_i\rangle|^2\delta\left(E_f-E_i-\hbar\omega_0\right)+O(\hat{W}_0^3) \tag{13.51}$$

Fermi's golden rule says that an excitation with monochromatic light is only possible if the incident light beam has exactly that energy $\hbar\omega_0$ that is necessary to lift the system into a higher-, or lower-energetic state.

With these transition rates we can investigate the **stimulated absorption** and **stimulated emission** of light. The stimulated emission of light underlies the function of a **laser**. Electrons are prepared in their excited state. They remain there until light stimulates emission, which pumps further energy into the laser beam. Important for the function of a laser is that the spontaneous emission of light is sufficiently small so that the system can be prepared in its excited state. The latter effect can be considered as the spontaneous emission of light induced by the zero point vibrations of the electromagnetic field[4].

It may be surprise the reader to see that the transition probability Eq. 13.51 depends quadratically rather than linearly on the perturbation. It makes sense if we consider that the energy density of the perturbing field depends quadratically on its amplitude. (See the energy density of the electromagnetic field.) Thus Fermi's golden rule is consistent with the expectation that a certain fraction of the incident light-energy is absorbed.

13.5 Variational or Rayleigh-Ritz principle

The variational principle is a useful scheme to find the ground state of a system. The **ground state** is the energy eigenstate with lowest energy. It is commonly known under the name Rayleigh-Ritz principle or Ritz-method.[41][10]

Nearly all methods to investigate the electrons in real solids or molecules are based on the variational principle.

The basis for the **variational principle** is the following observation:

- The energy $E[\psi] = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle}$ evaluated for an arbitrary state $|\psi\rangle$ is higher than the energy of the ground state, and
- that it this energy becomes equal to the ground-state energy, if we insert the ground state for $|\psi\rangle$

This is very handy if we cannot solve the Schrödinger equation exactly. We just choose an arbitrary function, and we can evaluate, if not the ground state energy itself, at least a strict upper bound for it. The better the guess for the wave function the closer we approach the ground state energy.

This is shown as follows:

[10]Walter Ritz, born 1878 in Sion, Switzerland, died 1909,Goettingen, Germany. swiss theoretical physicist, known for the variational principle named after him as well as for the Rydberg-Ritz formula, describing the spectral lines of the hydrogen atom. The contribution of Ritz was that the spectral lines can be constructed as differences of energy levels. Walter Ritz died in the young age of only 31 years on tuberculosis.

PROOF: let us expand $|\psi\rangle = \sum_n |\phi_n\rangle c_n$ with $c_n = \langle\phi_n|\psi\rangle$ into the orthonormal system of eigenstates $|\phi_n\rangle$ of the Hamiltonian with eigenvalues E_n, that is $\hat{H}|\phi_n\rangle = |\phi_n\rangle E_n$ and $\langle\phi_n|\phi_m\rangle = \delta_{n,m}$.

$$
\begin{aligned}
E[\psi] &\stackrel{\text{def}}{=} \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\langle\psi|\hat{H}\overbrace{\sum_n |\phi_n\rangle\langle\phi_n|}^{\hat{1}}|\psi\rangle}{\langle\psi|\underbrace{\sum_n |\phi_n\rangle\langle\phi_n|}_{\hat{1}}|\psi\rangle} \\
&= \frac{\sum_n \langle\psi\overbrace{|\phi_n\rangle E_n}^{\hat{H}|\phi_n\rangle}\overbrace{\langle\phi_n|\psi\rangle}^{c_n}}{\sum_n \langle\psi|\phi_n\rangle\langle\phi_n\psi\rangle} = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \\
E[\psi] - E_0 &= \frac{\sum_n |c_n|^2 (E_n - E_0)}{\sum_n |c_n|^2} \geq 0
\end{aligned}
$$

Since each term is non-negative the result must be non-negative. The equal sign holds, if all coefficients vanish except that for the ground state, i.e. for $E_n = E_0$. Hence the equal sign holds only for the ground state. *q.e.d*

The variational principle is the basis for almost all electronic structure methods, that try to determine the electronic wave functions for real materials. Here we work with N-particle systems, which are often approximated in one or the other way by non-interacting N-particle systems.

Since these systems are often enormously complex, we choose a set of basis functions $|\chi_i\rangle$. These functions could be atomic orbitals. In this case, the method would be called "linear superposition of atomic orbitals" (LCAO).[11]

Then we make an ansatz of these basis functions.

$$
|\psi_n\rangle = \sum_i |\chi_i\rangle c_{i,n}
$$

and minimize the total energy for this wave function.

If we multiply the Schrödinger equation from the left with a basis function, we obtain a eigenvalue equation in matrix form

$$
\begin{aligned}
\hat{H}|\psi_n\rangle &= |\psi_n\rangle E_n \\
\Rightarrow \sum_j \langle\chi_i|\hat{H}|\chi_j\rangle c_{j,n} &= \sum_j \langle\chi_i|\chi_j\rangle c_{j,n} E_n \\
\Rightarrow \boldsymbol{H}\vec{c}_n &= \boldsymbol{O}\vec{c}_n E_n
\end{aligned}
$$

where $H_{i,j} = \langle\chi_i|\hat{H}|\chi_j\rangle$ are the Hamilton matrix elements and $O_{i,j} = \langle\chi_i|\chi_j\rangle$ are the Overlap matrix elements. If the set of basis functions $|\chi_n\rangle$ is complete, this was nothing but a change of the representation. No approximation has been done so far.

However, if the basis set is complete, it is also infinite. Such a problem cannot be handled on the computer. In practice we thus have to limit the dimension of the vectors and matrices. Only a finite set of wave functions is used. This however is a fairly uncontrolled approximation. The variational principle provides us with two important conclusions about the approximate solution

- The energy is an upper bound of the ground state energy
- The deviations of the approximate wave functions only induces changes of second order in the total energy.

The groundstate of a non-interacting N-Fermion system is a Slater determinant formed by the N eigenstates of the corresponding one-particle Hamiltonian. The energy of this Slater determinant is

[11]Note that LCAO is not a very good technique for practical electronic structure methods.

given by the sum of the corresponding eigenvalues of the one-particle Hamiltonian. The variational principle shows that every Slater determinant of orthonormal orbitals has an energy larger or equal to the ground state.

$$
\begin{aligned}
E_0 &\le \min_{\vec{c}_0} F(\vec{c}_0, \lambda) \\
F(\vec{c}_0, \lambda) &= \vec{c}_0 \boldsymbol{H} \vec{c}_0 - \lambda\left(\vec{c}_0 \boldsymbol{O} \vec{c}_0 - 1\right) \\
\frac{\partial F}{\partial c_{i,0}} &= \sum_j H_{i,j} c_{j,0} - \lambda O
\end{aligned}
$$

where λ is a Lagrange multiplier for the constraint $\vec{c}_0 \boldsymbol{O} \vec{c}_0 = 1$, that the wave function is normalized.

PROOF: We describe the trial wave function $|\psi\rangle = \sum_n |\phi_n\rangle c_n$ as a linear superposition of basis functions $|\phi_n\rangle$ with variable coefficients c_n. Note, that the basis functions need not be complete, orthogonal or normalized.

$$
E[\psi] = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum_{n,m} c_n^* \langle\phi_n|\hat{H}|\phi_m\rangle c_m}{\sum_{n,m} c_n^* \langle\phi_n|\phi_m\rangle c_m}
$$

The minimum condition is $dE/dc_n = 0$ and $dE/dc_n^* = 0$. We have to consider here the coefficient and its complex conjugate as independent quantities. There are two derivatives, because each coefficient has a real and an imaginary part. The coordinate transformation from real and imaginary part to a complex coefficient and its complex conjugate is a common trick for handling complex numbers.

$$
\begin{aligned}
\frac{dE[\psi]}{dc_k^*} &= \frac{d}{dc_k^*}\Big(\frac{\sum_{n,m} c_n^* \langle\phi_n|\hat{H}|\phi_m\rangle c_m}{\sum_{n,m} c_n^* \langle\phi_n|\phi_m\rangle c_m}\Big) \\
&= \frac{\sum_m \langle\phi_k|\hat{H}|\phi_m\rangle c_m}{\sum_{n,m} c_n^* \langle\phi_n|\phi_m\rangle c_m} \\
&\quad - \frac{\sum_{n,m} c_n^* \langle\phi_n|\hat{H}|\phi_m\rangle c_m}{\sum_{n,m} c_n^* \langle\phi_n|\phi_m\rangle c_m} \times \frac{\sum_m \langle\phi_k|\phi_m\rangle c_m}{\sum_{n,m} c_n^* \langle\phi_n|\phi_m\rangle c_m}
\end{aligned}
$$

From the condition that the derivatives vanish, we obtain an equation for the coefficients

$$
\begin{aligned}
\Rightarrow 0 &= \sum_m \langle\phi_k|\hat{H}|\phi_m\rangle c_m - \sum_m \langle\phi_k|\phi_m\rangle c_m \bar{E} \\
\bar{E} &= \frac{\sum_{n,m} c_n^* \langle\phi_n|\hat{H}|\phi_m\rangle c_m}{\sum_{n,m} c_n^* \langle\phi_n|\phi_m\rangle c_m}
\end{aligned}
$$

What we obtained is a set of nonlinear equations for the coefficients, because $\bar{E}$ depends itself on the coefficients. However, the problem can be solved as eigenvalue problem, after identifying $\bar{E}$ with the lowest eigenvalue. *q.e.d*

The eigenvalue equation is nothing more than solving the Schrödinger equation in a limited Hilbert space defined by the basis functions. The more basis functions are included, the lower the energy and the more accurate is the result. By including more and more basis functions we can therefore approach the correct solution arbitrarily close.

The variational principle is so powerful, because any deviation in the wave function from the ground-state wave function enters only to second order into the energy expression.

PROOF: Here we show that the deviation of the trial wave function from the ground state wave function enters only to second order in the energy.
We expand the trial wave function $|\psi\rangle$ into eigenstates $|\phi_n\rangle$ of the Hamiltonian with energies E_n. The coefficients of the trial wave function $|\psi\rangle = \sum_n |\phi_n\rangle c_n$ are written as $c_n = \delta_{n,0} + \delta c_n$, where δc_n describes the deviation from the ground state $|\phi_0\rangle$.

$$
\begin{aligned}
E &= \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} \\
&= \frac{\sum_n |\delta_{n,0} + \delta c_n|^2 E_n}{\sum_n |\delta_{n,0} + \delta c_n|^2} = \frac{\sum_n (\delta_{n,0} + 2\mathrm{Re}[\delta c_n]\delta_{n,0} + |\delta c_n|^2) E_n}{\sum_n (\delta_{n,0} + 2\mathrm{Re}[\delta c_n]\delta_{n,0} + |\delta c_n|^2)} \\
&= \frac{E_0(1 + 2\mathrm{Re}[\delta c_0]) + \sum_n |\delta c_n|^2 E_n}{(1 + 2\mathrm{Re}[\delta c_0]) + \sum_n |\delta c_n|^2} = \frac{E_0 + \frac{\sum_n |\delta c_n|^2 E_n}{1+2\mathrm{Re}[\delta c_0]}}{1 + \frac{\sum_n |\delta c_n|^2}{1+2\mathrm{Re}[\delta c_0]}} \\
&= E_0 + \sum_n \frac{E_n - E_0}{1 + 2\mathrm{Re}[\delta c_0]} |\delta c_n|^2 + O(\delta c_n)^4 \\
&= E_0 + \sum_n (E_n - E_0)|\delta c_n|^2 + O(\delta c_n)^3
\end{aligned}
$$

In the last steps we used $\frac{1}{1+x} = 1 - x + O(x^2)$. Thus we find that $E = E_0 + O(\delta c_n^2)$ *q.e.d.*

A small error in the trial wave function produces an even smaller error in the energy. Hence, a very good estimate of the ground state energy can be obtained with fairly poor trial wave functions. It must be noted however, that the error in other observables is still of first order. One must not be misguided from a converged total energy, that the wave function itself is already of good quality. (A common mistake!)

Proof of the variational principle for non-interacting many-particle systems (Home study)

Let us start with N one particle states $|\psi_n\rangle$, which are orthonormal, that is $\langle\psi_n|\psi_m\rangle = \delta_{n,m}$. We need to show that

$$
\sum_{n=1}^{N} \langle\psi_n|\hat{H}|\psi_n\rangle = \sum_{n=1}^{N} E_n
$$

where E_n are the eigenvalues of the one-particle Hamiltonian. Let $|\phi_i\rangle$ be the exact eigenstates of the one-particle Hamiltonian.

$$
\begin{aligned}
\sum_{n=1}^{N} \langle\psi_n|\hat{H}|\psi_n\rangle &= \sum_{n=1}^{N} \langle\psi_n|\hat{H} \underbrace{\sum_j |\phi_j\rangle\langle\phi_j|}_{\hat{1}} |\psi_n\rangle = \sum_{n=1}^{N}\sum_j \langle\psi_n \overbrace{|\phi_j\rangle E_j}^{\hat{H}|\phi_j\rangle} \langle\phi_j|\psi_n\rangle \\
&= \sum_j E_j \left(\langle\phi_j| \underbrace{\sum_{n=1}^{N} |\psi_n\rangle\langle\psi_n|}_{=:\hat{P}} |\phi_j\rangle \right) = \sum_j E_j \langle\phi_j|\hat{P}_n|\phi_j\rangle
\end{aligned}
$$

We have introduced $\hat{P}$, the projection operator on the approximate wave functions chosen. We can now show that

- $0 \le \langle\phi_j|\hat{P}|\phi_j\rangle \le 1$.

- $\sum_j \langle\phi_j|\hat{P}|\phi_j\rangle = N$. This is shown as follows:

$$\sum_j \langle\phi_j|\hat{P}|\phi_j\rangle = \sum_j \langle\phi_j|\sum_{n=1}^{N}|\psi_n\rangle\langle\psi_n|\phi_j\rangle = \sum_{n=1}^{N}\langle\psi_n|\underbrace{\sum_j|\phi_j\rangle\langle\phi_j|}_{=\hat{1}}|\psi_n\rangle$$
$$= \sum_{n=1}^{N}\underbrace{\langle\psi_n|\psi_n\rangle}_{=1} = N$$

Now we can calculate the difference from the groundstate energy as

$$\begin{aligned}
E[\psi_1,\ldots,\psi_N] - \sum_{j=1}^{N} E_j \quad &= \quad \sum_{n=1}^{N}\langle\psi_n|\hat{H}|\psi_n\rangle - \sum_{n=1}^{N} E_n \\
&= \quad \sum_{j=1}^{N} E_j \underbrace{\langle\phi_j|\hat{P}-\hat{1}|\phi_j\rangle}_{\leq 0} + \sum_{j=N+1}^{\infty} E_j \underbrace{\langle\phi_j|\hat{P}|\phi_j\rangle}_{\geq 0} \\
&\overset{0\leq\langle\phi_j|\hat{P}|\phi_j\rangle\leq 1}{\geq} \quad \sum_{j=1}^{N} E_N \underbrace{\langle\phi_j|\hat{P}-\hat{1}|\phi_j\rangle}_{\leq 0} + \sum_{j=N+1}^{\infty} E_{N+1} \underbrace{\langle\phi_j|\hat{P}|\phi_j\rangle}_{\geq 0} \\
&= \quad (E_{N+1}-E_N)\sum_{j=N+1}^{\infty}\langle\phi_j|\hat{P}|\phi_j\rangle \geq 0
\end{aligned}$$

13.6 WKB-Approximation

The WKB approximation is useful for quasi-one-dimensional potentials, which have little structure on a length scale of the wave length of the particle considered. It is used for example in semiconductor physics to investigate the transport of charge carriers.

A generalization to time-dependent problems and to higher dimensions is provided in appendix K.1on p. 303.

The **WKB approximation**[12] named after Wentzel, Kramers and Brillouin[42, 43, 44] is also called the **quasi-classical approximation**.

A very good description of he WKB approximation can be found in the book by Griffiths[5].

Apparently, the essential features of the WKB approximation have been known as early as 1912[45]. **The WKB method is applicable if the potential is rather smooth, so that is it varies only little over a wave length of the particle.** It is also called quasi-classical approximation, because it is a type of perturbation theory, where the small parameter is Planck's constant $\hbar$. (Recall that $\hbar = 0$ is the classical limit.)

For a given potential, the wavelength of a particle with a given energy scales with $\hbar$. Thus the smaller $\hbar$, the better the condition, that the potential varies slowly over the wavelength of a particle, is fulfilled. Even though the WKB approximation can also be applied in higher dimensions, it most commonly applied only in one dimension, where it is particularly simple. The generalization of the WKB method to higher dimensions is nontrivial. It is described in appendix K.1 on p. 303

[12] In the Netherlands it is called KWB, in France BWK and in Britain JWKB (for Jeffreys). Kramers was dutch, Brillouin french and Jeffreys was british.

The WKB method starts with the following Ansatz for the wave function

$$
\begin{aligned}
\psi(x) &= \mathrm{e}^{\frac{i}{\hbar}S(x)} \\
\Rightarrow \hat{p}\psi(x) &= \frac{\hbar}{i}\partial_x \mathrm{e}^{i\frac{S(x)}{\hbar}} = \mathrm{e}^{i\frac{S(x)}{\hbar}}\Big(\partial_x S\Big) \\
\Rightarrow \hat{p}^2\psi(x) &= [\frac{\hbar}{i}\partial_x]^2 \mathrm{e}^{i\frac{S(x)}{\hbar}} = \mathrm{e}^{i\frac{S(x)}{\hbar}}\Big((\partial_x S)^2 + \frac{\hbar}{i}\partial_x^2 S\Big)
\end{aligned}
$$

When we insert this result into the one-dimensional Schrödinger equation

$$
E\psi = \Big[\frac{-\hbar^2}{2m}\partial_x^2 + V(x)\Big]\psi \tag{13.52}
$$

we obtain

$$
E = \frac{1}{2m}(\partial_x S)^2 + V(x) + \frac{\hbar}{2mi}(\partial_x^2 S) \tag{13.53}
$$

Note that we keep the energy fixed in this case. The typical application is to find an approximation for a wave function at a given energy irrespective of the boundary conditions.[13]

Now, we expand S in powers of $\hbar$, that is

$$
S = S_0(x) + \hbar S_1(x) + \ldots
$$

We insert this series into the equation (13.53) and order the terms according to powers of $\hbar$.[14]

$$
\begin{aligned}
0 &= \frac{1}{2m}(\partial_x S_0 + \hbar\partial_x S_1 + \ldots)^2 + V(x) - E + \frac{\hbar}{2mi}(\partial^2 S_0 + \hbar\partial_2 S_1 + \ldots) \\
&= \Big[\frac{1}{2m}(\partial_x S_0)^2 + V(x) - E\Big] \\
&\quad + \Big[\frac{1}{m}(\partial_x S_0)(\partial_x S_1) + \frac{1}{2mi}\partial_x^2 S_0\Big]\hbar + O(\hbar^2)
\end{aligned} \tag{13.54}
$$

For the lowest order, we obtain an expression for S_0

$$
\frac{(\partial_x S_0)^2}{2m} + V(x) - E = 0
$$

This is nothing but the classical energy momentum relation, if we identify S_0 with the classical action and $\partial_x S_0$ with the momentum.[15]

[13] This differs from the applications of first order perturbation theory, where we wanted to find a better bound state. Thus the boundary conditions have been fixed, but the energy depended on the perturbation.

[14] Note that this step is analogous to that in first order perturbation theory.

[15] The momentum is indeed the gradient of the action which can be shown as follows. We start with the variation of the action. However, in contrast to the derivation of the Euler Lagrange equations, we allow also variations of the end points of the trajectory.

$$
\begin{aligned}
S[x(t)] &= \int dt\ \mathcal{L}(x(t), \dot{x}(t), t) \\
\delta\mathcal{S} = S[x(t)+\delta x(t)] - S[x(t)] &= \int dt\ \Big[\frac{\partial\mathcal{L}}{\partial x}\delta x(t) + \frac{\partial\mathcal{L}}{\partial\dot{x}}\delta\dot{x}(t)\Big] \\
&= \int dt\ \Big[\frac{\partial\mathcal{L}}{\partial x}\delta x(t) + \frac{d}{dt}\Big(\frac{\partial\mathcal{L}}{\partial\dot{x}}\delta x(t)\Big) - \delta x(t)\frac{d}{dt}\Big(\frac{\partial\mathcal{L}}{\partial\dot{x}}\Big)\Big] \\
&= \int dt\ \delta x(t)\underbrace{\Big[\frac{\partial\mathcal{L}}{\partial x}\delta x(t) - \frac{d}{dt}\Big(\frac{\partial\mathcal{L}}{\partial\dot{x}}\Big)\Big]}_{ELG} + \int dt\ \frac{d}{dt}\Big(\frac{\partial\mathcal{L}}{\partial\dot{x}}\delta x(t)\Big)
\end{aligned}
$$

If we consider now physical trajectories, which fulfill the Euler-Lagrange equation, the term denoted by "ELG" vanishes

$$
\begin{aligned}
p(x) &= \partial_x S_0 = \pm\sqrt{2m(E-V(x))} \qquad (13.55)\\
S_0(x) &= \pm\int^x dx' p(x') + \text{const.} = \pm\int^x dx' \sqrt{2m(E-V(x'))} + \text{const.}
\end{aligned}
$$

We obtained a closed expression for the lowest term. The corresponding wave function is oscillatory, but with a local wave vector $\hbar k = \frac{dS}{dx}$ that depends on space.

$$
\psi(x) = A\mathrm{e}^{\frac{i}{\hbar}S_0(x)} + B\mathrm{e}^{-\frac{i}{\hbar}S_0(x)}
$$

In the classically allowed region, i.e. where $V(x) < E$, the maxima and minima of the wave function remain constant in space.

In the classically forbidden region, the momentum is imaginary. $p(x) = \pm i\lambda(x)$ with $\lambda(x) = \sqrt{2m(V-E)}$. is imaginary. Therefore, also the action is imaginary, so that the wave function consist out of an exponentially decreasing and an exponentially increasing contribution.

Let us now consider the next higher order in Eq. 13.54.

$$
\begin{aligned}
0 &\stackrel{\text{Eq. 13.54}}{=} \frac{1}{m}(\partial_x S_0)(\partial_x S_1) + \frac{1}{2mi}\partial_x^2 S_0\\
0 &= p(x)(\partial_x S_1) - \frac{i}{2}\partial_x p(x)\\
\partial_x S_1 &= \frac{i}{2}\frac{\partial_x p(x)}{p(x)} = \frac{i}{2}\partial_x \ln[|p(x)|] = \partial_x i\ln[\sqrt{|p(x)|}]\\
S_1 &= i\ln[\sqrt{|p(x)|}] + \text{const.}
\end{aligned}
$$

Thus we obtain for as result for the wave function which now includes also an amplitude modulation.

WAVE FUNCTION IN THE WKB APPROXIMATION

$$
\psi(x) = \frac{1}{\sqrt{|p(x)|}}\left[A\mathrm{e}^{\frac{i}{\hbar}\int dx p(x)} + B\mathrm{e}^{-\frac{i}{\hbar}\int dx p(x)}\right]
$$

where the momentum $p(x)$ is according to Eq. 13.55

$$
p(x) = \sqrt{2m(E-V(x))}
$$

The WKB approximation always requires the first two terms, S_0 and S_1, to provide a meaningful result. This is because there is a prefactor $1/\hbar$ in front of the expansion. Thus, the second term, S_1, is required to obtain the wave function to zeroth order in $\hbar$.

The zeroth order, however, plays an important historical role, because it captures **Bohr's quantization condition**, which was one of the first steps towards quantum theory. In that theory, the energy levels have been quantized by the requirement that only those closed classical trajectories

and only the last term remains. Noting that the canonical momentum is defined as $p = \partial\mathcal{L}/\partial\dot{x}$, we obtain

$$
\delta S = \int_{t_i}^{t_f} dt\, \frac{d}{dt}\left(\underbrace{\frac{\partial\mathcal{L}}{\partial\dot{x}}}_{p}\delta x(t)\right) = p(t_f)\delta x(t_f) - p(t_i)\delta x(t_i)
$$

If we now consider the variation of the action under a variation of the end point along the trajectory, we obtain

$$
\frac{dS}{dx(t_f)} = p(t_f)
$$

are allowed, for which the change of the phase $S_0(x)$ along a closed classical path is a multiple of Planck's quantum $h = 2\pi\hbar$, that is $\int dx p(x) = \hbar n$. This empirical assumption lead to **Bohr's atom model**, which captured the observed excitation energies between the states of the hydrogen atom. Note that along a closed trajectory S_0 must change by $2\pi n$ along a closed trajectory in order to avoid a discontinuity in the wave function.

13.6.1 Classical turning points

We can already see that the wave function diverges at classical turning points, that is the points where $E = V(x)$. At those points the series expansion Eq. 13.54 does not converge.

Let us investigate Eq. 13.53 and resolve it for $\partial_x S$

$$(\partial_x S) = \pm\sqrt{2m(E - V(x)) - \frac{\hbar}{i}(\partial_x^2 S)} \tag{13.56}$$

$$= \pm\sqrt{2m(E - V(x))}\sqrt{1 + i\frac{\hbar(\partial_x^2 S)}{2m(V - E)}} \tag{13.57}$$

The expansion of S in orders of $\hbar$ is identical to a Taylor expansion of the last factor. Since the convergence radius of the Taylor series of $\sqrt{1+x}$ is equal to one[16] , the expansion converges only if

$$|\partial_x^2 S| < \frac{2m|E - V|}{\hbar}$$

At the classical turning point with $V(x) = E$, this condition is violated, unless $\partial_x^2 S$ vanishes to all orders in $\hbar$.

Near classical turning points we need to evaluate the Schrödinger equation directly. The problem can be simplified if the potential is linearized near the turning point. The comparison of asymptotic solutions of the exact solution and that of the WKB approximation provides connection formulas. We will not go into that detail, but only state the final result. The derivation is given in the appendix App. K.2. Here we only list the final result.

MATCHING CONDITION FOR THE WKB SOLUTION AT THE CLASSICAL TURNING POINT

The WKB wave function for a potential with positive slope at the classical turning point can be written as

$$\Psi(x < x_{cl}) = \frac{1}{\sqrt{|p(x)|}}\cos(\frac{1}{\hbar}[\int_x^{x_{cl}} dx' p(x')] - \frac{\pi}{4})$$

$$\Psi(x > x_{cl}) = \frac{1}{\sqrt{|p(x)|}}\mathrm{e}^{-\frac{1}{\hbar}\int dx' p(x')}$$

and for negative slope

$$\Psi(x > x_{cl}) = \frac{1}{\sqrt{|p(x)|}}\cos(\frac{1}{\hbar}[\int_{x_{cl}}^{x} dx' p(x')] - \frac{\pi}{4})$$

$$\Psi(x < x_{cl}) = \frac{1}{\sqrt{|p(x)|}}\mathrm{e}^{\frac{1}{\hbar}\int dx' p(x')}$$

[16] The convergence radius of a Taylor expansion is at most the distance of the nearest singularity. The square root has a diverging derivative at the origin, that is for $x_{sing} = -1$ in $\sqrt{1+x}$. Thus the convergence radius is at most $|x_{sing}| = 1$

13.7 Numerical integration of one-dimensional problems

The one-dimensional Schrödinger equation can readily be integrated numerically.

$$[\frac{-\hbar^2}{2m}\partial_x^2+V]\psi=E\psi$$

We use here a method, that is completely analogous to the Verlet, which has been introduced in ΦSX: Klassische Mechanik[14] to numerically solve Newton's equations of motion. The basic idea is to place the wave function onto a grid $x_i=\Delta i$, where Δ is the grid spacing and i is the integer index of the grid point. Now we replace the derivative by the corresponding differential quotients

$$\partial_x\psi(x_i)\approx\frac{\psi_{i+1}-\psi_{i-1}}{2\Delta}$$
$$\partial_x^2\psi(x_i)\approx\frac{\psi_{i+1}-2\psi_i+\psi_{i-1}}{\Delta^2}$$

and obtain the discretized Schrödinger equation[17]. We resolve it for ψ_{i+1}

$$\frac{-\hbar^2}{2m\Delta^2}(\psi_{i+1}-2\psi_i+\psi_{i-1})+V\psi_i=E\psi_i$$
$$\psi_{i+1}=2\psi_i-\psi_{i-1}-\frac{2m(E-V(x_i))\Delta^2}{\hbar^2}\psi_i$$

Thus knowing the value of ψ at two successive grid points we can evaluate the wave function recursively on all other grid points. With the formula given above we obtain the wave function for increasing x. If we resolve the discretized Schrödinger equation for ψ_{i-1}, we obtain a recursive equation that leads towards smaller x.

13.7.1 Stability

Numerical solutions are not exact. Two aspects need to be understood. The first is accuracy and the second is stability. Accuracy can be improved by decreasing the grid spacing Δ. The computation becomes more involved so that we have to find a good compromise between accuracy and computational effort. When we try to minimize the computational effort by increasing the grid spacing we run into the problem that the scheme may become unstable, and that it will produce qualitatively incorrect results.

Consider for example the classically forbidden region, where we have exponentially increasing and exponentially decreasing solutions. If we integrate downward an exponentially decreasing solution, a small error will create a small component of the increasing solution, and therefore magnify. Such a procedure is therefore numerically unstable. If, on the other hand, we integrate in the opposite direction, and integrate the solution in choosing the direction in which it increases, any error will exponentially die out. Therefore, one is not entirely free in the direction: In the classical forbidden region one should always propagate the solution in the direction where the wave function increases in absolute value.

[17] The accuracy of these expressions can be tested by inserting the Taylor expansion of ψ

$$\psi_{i+1}=\psi_i+\Delta\partial_x\psi+\frac{1}{2}\Delta^2\partial_x^2\psi++\frac{1}{6}\Delta^3\partial_x^3\psi+\frac{1}{24}\Delta^4\partial_x^4\psi+\dots$$
$$\psi_{i-1}=\psi_i-\Delta\partial_x\psi+\frac{1}{2}\Delta^2\partial_x^2\psi+-\frac{1}{6}\Delta^3\partial_x^3\psi+\frac{1}{24}\Delta^4\partial_x^4\psi+\dots$$
$$\frac{\psi_{i+1}-\psi_{i-1}}{2\Delta}=\partial_x\psi+\frac{1}{6}\Delta^2\partial_x^3\psi+O(\Delta)^4$$
$$\frac{\psi_{i+1}-2\psi_i+\psi_{i-1}}{\Delta^2}=\partial_x^2\psi+\frac{1}{12}\Delta^2\partial_x^4\psi+O(\Delta)^4$$

Thus the error introduced is of order Δ^2

If one follows an exponentially increasing wave function we need to remember that a computer can only represent numbers up to a maximum value, which is 1.79×10^{308} for an 8 byte real number. While this is a large number, it can still be reached quickly with an exponential increasing function. It is therefore necessary to check the size of the wave function and rescale the entire solution (on all grid points). A computer is not only limited in representing large numbers, but there is also a smallest number the computer can describe. Where the absolute value of the wave function drops below 2×10^{-308}, the smallest number on an 8-Byte computer, the wave function will be set exactly to zero, but there the result can be considered insignificant for most purposes.

For an oscillatory solution we need to choose the step size sufficiently small. One should choose the step size smaller than one percent of the period of the oscillation, that is

$$\Delta < \frac{1}{10}\frac{2\pi\hbar}{\sqrt{2m(E-\min[V])}}$$

If the step size is larger than one third of the oscillation period, the algorithm becomes unstable.

13.7.2 Natural boundary conditions

For bound states an important question is how to start up the solution. We cannot use two zero values of the wave function because then the result will be the zero state. Here we can use the WKB approximation. We use the function

$\psi(x) = \frac{1}{\sqrt[4]{2m(V(x)-E)}} \mathrm{e}^{-\frac{1}{\hbar}\int dx\sqrt{2m(V(x)-E)}}$

to evaluate the wave function on a grid starting from a point where the WKB approximation is safe. Then we use two such values to follow the solution into the classically allowed region.

We can do this from the left and from the right, but since we do not yet know the correct energy, the solution will not match, where they meet. Here we can use another trick.

We expand the solution in a Taylor series $\psi(E) = \psi(E_0) + (E - E_0)\frac{d\psi}{dE}$ about one energy. To first order

$$[\frac{-\hbar^2}{2m}\partial_x^2 + V - E]\frac{d\psi}{dE} = \psi(E)$$

We discretize this differential equation to obtain $\frac{d\psi}{dE}$. To start the solution up, we must of course also form the energy derivative of the WKB solution. Then we decide on a matching point in the classically allowed region, and evaluate values and derivatives using the differential quotient.

A prediction of the energy can be obtained from the condition that the logarithmic derivatives match

$$\frac{\partial_x\psi_+ + dE\partial_x\dot{\psi}_+}{\psi_+ + dE\dot{\psi}_+} = \frac{\partial_x\psi_- + dE\partial_x\dot{\psi}_-}{\psi_- + dE\dot{\psi}_-}$$

which results in a quadratic equation for $dE = E - E_0$

$$A + dEB + CdE^2 = 0$$
$$A = \psi_-\partial_x\psi_+ - \psi_+\partial_x\psi_-$$
$$B = \dot{\psi}_-\partial_x\psi_+ - \dot{\psi}_+\partial_x\psi_- + \psi_-\partial_x\dot{\psi}_+ - \psi_+\partial_x\dot{\psi}_-$$
$$C = \dot{\psi}_-\partial_x\dot{\psi}_+ - \dot{\psi}_+\partial_x\dot{\psi}_-$$

We first test if there is a solution by checking if $B^2 - 4AC > 0$. If there is no such solution, we linearize the equation and use $E = E_0 - B/A$ as the next better estimate for the energy.

If there is such a solution for dE from the quadratic equation, we solve this equation and take the solution for dE with the smaller absolute value to obtain a new estimate for the energy $E = E_0 + dE$. Using this value, we repeat the procedure until the energy is converged. Now we rescale the solutions so that their values are identical at the matching points.

13.7.3 Radial wave functions for atoms

The radial Schrödinger equation has the form $\psi(\vec{r}) = \phi_{\ell,m}(|r|)Y_{\ell,m}(\vec{r})$

$$\begin{aligned}
0 &= \Big[\frac{-\hbar^2}{2m}\frac{1}{r}\partial_r^2 r + \frac{\hbar^2\ell(\ell+1)}{2mr^2} + V(r) - E\Big]\phi_{\ell,m}(r) \\
&= \frac{1}{r}\Big[\frac{-\hbar^2}{2m}\partial_r^2 + \frac{\hbar^2\ell(\ell+1)}{2mr^2} + V(r) - E\Big] r\phi_{\ell,m}(r)
\end{aligned}$$

For an atom the potential has singularities at the origin. Therefore, a uniform step-size will not be appropriate. We need to choose a step size that becomes finer near the origin. Therefore, we introduce a logarithmic grid. First we introduce a variable transform

$$r(x) = \mathrm{e}^{\alpha x}$$

which results in

$$\begin{aligned}
\frac{df}{dx} &= \frac{df}{dr}\frac{dr}{dx} = \alpha r\frac{df}{dr} \Rightarrow \partial_r = \frac{1}{\alpha r}\partial_x \\
\partial_r^2 &= -\frac{1}{\alpha r^2}\partial_x + \frac{1}{\alpha^2 r^2}\partial_x^2
\end{aligned}$$

and obtain, after multiplication from the left with r, the transformed radial Schrödinger equation

$$\begin{aligned}
&\Big[\frac{-\hbar^2}{2m\alpha^2 r^2}\partial_x^2 - \frac{-\hbar^2}{2m\alpha r^2}\partial_x + \frac{\hbar^2\ell(\ell+1)}{2mr^2} + V(r) - E\Big] r\phi_{\ell,m}(r) = 0 \\
\Rightarrow &\Big[\frac{-\hbar^2}{2m}(\partial_x^2 - \alpha\partial_x) + \frac{\alpha^2\hbar^2\ell(\ell+1)}{2m} + \alpha^2 r^2(V(r) - E)\Big] r\phi_{\ell,m}(x) = 0
\end{aligned}$$

This form has no more divergent terms at the origin. We discretize along x and use standard procedures. The grid becomes infinitely fine at the origin, so that the divergent terms can be easily integrated. At the origin of r that is for $x \to -\infty$ we need to obey the correct initial conditions. We choose the innermost grid point and use the asymptotic solution at the origin to evaluate the wave function at the first two grid points.

13.8 Recommended exercises

1. Optical absorption: Exercise 16.8 on p. 252. (Not ready)
2. Alpha decay: Exercise 16.9 on p. 255. (Not ready)
3. Tunneling barrier. Exercise 16.10 on p. 255. (Not ready)
4. Transmission coefficient. Exercise 16.11 on p. 255. (Not ready)

Chapter 14

Relativistic particles

In the first chapter we have learned that the hanging linear chain describes relativistic particles, and that the Schrödinger equation is the non-relativistic approximation. While the Schrödinger equation is appropriate to describe most observations, there are instances where relativistic effects are important. For example, close to the nucleus, electrons gain sufficient kinetic energy that their velocity becomes comparable to the speed of light. Therefore, atoms contain relativistic effects from that region.

14.1 A brief review of theory of relativity

We are already familiar with the Klein-Gordon equation which we derived from the linear chain after introducing a mass term. After applying the correspondence principle, the Klein-Gordon equation resulted in the relativistic dispersion relation $E(p)$ with $E^2 = p^2c^2 + m_0^2c^4$. All relativistic particles obey the same dispersion relation. This dispersion relation lead Einstein to propose the famous equivalence of energy and mass $E = m_0c^2$, which is at the root of nuclear power, peaceful or deadly.

This dispersion relation is symmetric with respect to transformations that leave $E^2 - p^2c^2$ invariant. Energy and momentum are related through the correspondence principle to derivatives of space and time. This implies that $\frac{1}{c^2}\partial_t^2 - \nabla^2$ is an invariant of the symmetry groups of free space. It might be interesting to find out what these symmetries are and if we can learn something from the symmetry of the dispersion relation about the geometric structure of space and time.

The relativistic dispersion relation can be written with the help of the metric tensor[1]

$$g_{i,j} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and the four-dimensional momentum $p = (E/c, p_x, p_y, p_z) = \frac{\hbar}{i}(-\frac{1}{c}\partial_t, \nabla)$ as

$$\sum_{i,j=0}^{3} (p_i c) g_{i,j}^{-1} (p_j c) + m_0^2 c^4 = 0$$

We use here the notation that the momentum has a zeroth component, which is related to the energy by $p_0 = E/c$. The momentum is a four dimensional vector just as time and space are written in the formulation of the theory of relativity as four dimensional vector (ct, x, y, z). The components

[1] There are a number of different notations that are equivalent. Often the element -1 of the metric tensor is written as $-\frac{1}{c^2}$ and often the metric tensor has opposite sign. These are all equivalent notations. I have chosen one where the metric tensor is dimensionless, and where the four-dimensional momentum vector has dimensions mass*length/time

of the four dimensional momentum vector are the canonical conjugate momenta of the corresponding four-dimensional coordinate vector. For example the energy is conserved if the equations of motion are translationally invariant in time. The momentum is conserved if they are translational invariant in space.

Let us now consider coordinate transformations that leave $\frac{1}{c^2}\partial_t^2 - \nabla^2$ invariant. Let us write the coordinate transformation as

$$x'_i = \sum_{j=0}^{4} \frac{\partial x'_i}{\partial x_j} x_j = \sum_{j=0}^{4} R_{i,j} x_j$$

What is the corresponding transformation of derivatives?

$$\partial_i \psi = \sum_{j=0}^{4} \frac{\partial x'_j}{\partial x_i} (\partial'_j \psi) = \sum_{j=0}^{4} (R_{j,i} \partial'_j) \psi$$

We have used here [2] the vector $\partial_i = \frac{\partial}{\partial x_i}$.

In a vector notation we can write

$$x' = Rx$$
$$\partial = R^T \partial'$$

The symmetry operations of free space are coordinate transformations R that leave $\partial g^{-1} \partial = -\frac{1}{c^2}\partial_t^2 + \nabla^2$ invariant. These operations are called **Lorentz transformations** What are those transformations?

$$\partial g^{-1} \partial = \partial' R g^{-1} R^T \partial' = \partial' g^{-1} \partial'$$
$$R g^{-1} R^T = g^{-1}$$

This is the defining equation for the symmetry operations.

We can show that all those transformations leave a certain distance $d = \sqrt{c^2t^2 - r^2}$ in the four-dimensional space invariant.

$$x' g x' = x R^T g R x = x(gR^{-1}g^{-1}) g R x = x g x$$
$$c^2 t'^2 - r'^2 = c^2 t^2 - r^2$$

The invariance of this length is taken to define the basic symmetry of space. What is the distance between two space-time points. Two points that are spatially distant can still have zero distance in space time, namely when it takes the speed of light to get from (r_1, t_1) to (r_2, t_2).

Now we can investigate the symmetry operations in more detail: Let us consider a few examples. Translations in time and space conserve d. Also Rotations conserve the length. The rotation in the x,y-plane can be written as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos(\phi_z) & \sin(\phi_z) \\ -\sin(\phi_z) & \cos(\phi_z) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

When we replace x in the four-dimensional distance by x', we obtain the same result. Similarly if a physical path of a particle is $x(t)$ then also $x'(t')$ is a physical path, that obeys the equations of motion.

[2] Here a more detailed derivation:

$$\frac{\partial \Psi(x'_1, x'_2, x'_3, x'_4)}{\partial x_i} = \sum_j \frac{\partial \Psi(x'_1, x'_2, x'_3, x'_4)}{\partial x'_j} \frac{\partial x'_j(x_1, x_2, x_3, x_4)}{\partial x_i} = \sum_j \frac{\partial \Psi}{\partial x'_j} R_{j,i} = \sum_j R^T_{i,j} \frac{\partial \Psi}{\partial x'_j}$$

There must be also a similar transformation in the ct, x plane that conserves the four-dimensional distance

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh(\frac{v}{c}) & \sinh(\frac{v}{c}) \\ \sinh(\frac{v}{c}) & \cosh(\frac{v}{c}) \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}$$

It can be easily shown using $\cosh^2(x) - \sinh^2(x) = 1$ that the distance $c^2t^2 - x^2$ is conserved. What is the physical meaning of this transformation? Let us investigate the nonrelativistic limit, that is $c \to \infty$. We obtain the Galileo transformation

$$x' = x - vt$$
$$t' = t$$

Invariance with respect to this transformation simply means that if a path physical $x(t)$ obeys the equations of motion, also the path $x'(t') = x(t) + vt$ obeys the same equations of motion. This is of course true for the free particle. The transformed path is the original $x(t)$ as it is seen by an observer moving with velocity $-v$. If the speed of light c is finite, as it is in reality, the transformation of the equations of motion is more complex.

In order to build a relativistic theory, we need only to start combining so-called **invariants** of the symmetry group to build up the Lagrangian, from which the equations of motion are derived. Once we have such a theory, we might look for particles in the real world that are described by such a theory. This is one way to come up with theories for new particles.

We have already seen that $E^2 - p^2c^2$ is such an invariant and another is $m_0^2c^4$. Adding the two invariants, we arrive at the relativistic dispersion relation, from which we started. The correspondence principle brings us back to the Klein-Gordon equation, which describes so-called scalar particles. However there are other invariants, for example we might express the wave function as an array of wave functions in each point. The wave function of such a particle is called a spinor wave function. Now we can build up new invariants from operators acting on the spinor components, which still are Lorentz invariant in space-time. One very important vector particle are photons, which are spin-one particles. Another are electrons and other particles with spin $\frac{1}{2}$.

14.2 Relativistic Electrons

The Dirac equation can be written as

$$\Big(E - \sum_i \alpha_i c p_i - \beta m_0 c^2\Big)|\Psi\rangle = 0 \tag{14.1}$$

where the energy is here an operator, namely $E = i\hbar\partial_t$ and the momentum is the momentum operator $p = \frac{\hbar}{i}\nabla$. The factors α_i and β_i are themselves operators acting not on the spatial variable but in a four-dimensional spinor space. We will come to that later. For now we need to remember that α_i and β may not commutate among each other.

One condition for a relativistic wave equation is that it reproduces the relativistic dispersion relation $E^2 = p^2c^2 + m_0^2c^4$. We get there by multiplying the Dirac equation from the left with $E + \sum_i \alpha_i c p_i + \beta m_0 c^2$.

$$\Big(E^2 - \sum_{i,j=1}^{3} \alpha_i p_i c^2 \alpha_j p_j - \sum_{i=1}^{3} m_0 c^3 [\alpha_i \beta + \beta \alpha_i] p_i - m_0^2 c^4 \beta^2\Big)|\Psi\rangle = 0$$

The relativistic dispersion operator is obtained when the conditions

$$\frac{1}{2}[\alpha_i, \alpha_j]_+ = \delta_{i,j} \qquad \frac{1}{2}[\alpha_i, \beta]_+ = 0 \qquad \frac{1}{2}[\beta, \beta]_+ = 1$$

These equations cannot be fulfilled when α and β are simple numbers. A solution can be found by assuming that the α_i and β_i are matrices in a four dimensional space. This statement does not refer to the space-time, but to four components of the wave function $|\Psi\rangle$.

There is one solution, namely

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \qquad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

where the σ_i's are the Pauli matrices and the 1's are two-by-two unit matrices.

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The four spinor components of the wave function $|\Psi\rangle$ describe

- an electron with spin up
- an electron with spin down
- a positron with spin up
- a positron with spin down

Thus the **Dirac equation**, which describes leptons such as electrons, for free particles has the form

$$\Big[i\hbar\partial_t \mathbf{1} - \frac{\hbar c}{i}\sum_{j=1}^{3} \boldsymbol{\alpha}_j \partial_{x_i} - m_0 c^2 \boldsymbol{\beta}\Big]|\Psi\rangle = 0 \tag{14.2}$$

Note that the wave function $|\Psi\rangle$ is four component spinor, and that the matrices $\mathbf{1}$, $\boldsymbol{\alpha}_j$, and $\boldsymbol{\beta}$ are matrices acting on this four component spinor.

Positrons[46] are antiparticles of the electron. They have identical mass but opposite charge as the electron. A positron and an electron can annihilate each other by creating two photons. The photons carry the energy of the rest mass away.

14.3 Electron in the electromagnetic field

The electron(positron) is coupled to the electromagnetic field by replacing

$$E \to E - q\Phi$$
$$\vec{p} \to \vec{p} - q\vec{A}$$

where $\Phi(r)$ is the electric potential and $\vec{A}$ is the vector potential. They are defined as $\vec{E} = -\vec{\nabla}\phi - \partial_t\vec{A}$ and $\vec{B} = \vec{\nabla} \times \vec{A}$.The charge of the electron is negative, that is $q = -e$ where e is the elementary charge, which is positive.

Thus the Dirac equation has the form

$$\Big(E - q\Phi - \sum_i \alpha_i c(p_i - qA_i) - \beta m_0 c^2\Big)|\Psi\rangle = 0 \tag{14.3}$$

If we split the four-component spinor into two two-component spinors ($|\phi\rangle, |\chi\rangle$), the Dirac equation has the form

$$\begin{pmatrix} m_0c^2 + q\Phi - E & \sum_i \sigma_i(p_i - qA_i)c \\ \sum_i \sigma_i(p_i - qA_i)c & -m_0c^2 + q\Phi - E \end{pmatrix} \begin{pmatrix} |\phi\rangle \\ |\chi\rangle \end{pmatrix} = 0 \tag{14.4}$$

where the first two component spinor $|\phi\rangle$ is attributed to electronic degrees of freedom for spin-up and spin-down electrons and $|\chi\rangle$ describes spin-up and spin-down positron.

14.4 Down-folding the positron component

The four components of the Dirac spinor can be attributed to electrons with spin up and spin down and positrons with spin up and spin down. The positron is the anti-particle of the electron. While a Dirac particle is usually a superposition of electronic and positronic components, in the non-relativistic limit, the two different components can be decoupled exactly.

In this case, we are interested in describing electrons. Therefore, we consider solution with an energy close to $E = +m_0c^2$. In this case, we can show that the positronic components are fairly small. Therefore, one often names the electronic components $|\phi\rangle$ the "large components", while the positronic components $|\chi\rangle$ are named "small component". We do not want to ignore the positronic components, but we can simplify the equations by eliminating the positronic components in an effective way.

In order to simplify the following equations let us introduce a quantity

$$M(\epsilon, r) = m_0 + \frac{\epsilon - q\Phi}{2c^2}$$

and the energy ϵ with the rest-energy removed

$$\epsilon = E - m_0c^2$$

Note that M is not the relativistic mass defined as $m_r = m_0\sqrt{1 + (\frac{p}{m_0c})^2} = m_0 + \frac{p^2}{2m_0c^2} + O(\frac{1}{c^2}) = m_0 + \frac{\epsilon - V}{c^2}$.

Thus we obtain

$$\begin{pmatrix} q\Phi - \epsilon & \sum_i \sigma_i(p_i - qA_i)c \\ \sum_i \sigma_i(p_i - qA_i)c & -2Mc^2 \end{pmatrix} \begin{pmatrix} |\phi\rangle \\ |\chi\rangle \end{pmatrix} = 0$$

We can now use the second line of the equation to obtain $|\chi\rangle$ as function of the electron components $|\phi\rangle$,

$$|\chi\rangle = \frac{1}{2Mc}\sum_i \sigma_i(p_i - qA_i)|\phi\rangle,$$

which we then insert in to the first equation to eliminate the explicit dependence on the positrons.

$$\Big(\sum_i \sigma_i(p_i - qA_i)\frac{1}{2M}\sum_j \sigma_j(p_j - qA_j) + q\Phi - \epsilon\Big)|\phi\rangle = 0$$

We can use the equation

$$(\vec{\sigma}\vec{A})(\vec{\sigma}\vec{B}) = \vec{A}\vec{B} + i\vec{\sigma}(\vec{A} \times \vec{B}) \tag{14.5}$$

which connects the vectors $\vec{A}$ and $\vec{B}$ and the vector formed by the Pauli matrices. When we apply this equation to operators one has to take care of the commutation relations.

$$\Big[\sum_i (p_i - qA_i)\frac{1}{2M}(p_i - qA_i) + q\Phi + i\sum_i \sigma_i\Big(\vec{p} - q\vec{A}\big) \times \frac{1}{2M}(\vec{p} - q\vec{A})\Big)_i - \epsilon\Big]|\phi\rangle = 0 \tag{14.6}$$

Note that this equation is still nonlinear in energy and electromagnetic field, via the dependence on the relativistic mass.

Furthermore the wave function ϕ is not normalized. The correct normalization is

$$1 = \langle\phi|\phi\rangle + \langle\chi|\chi\rangle = \langle\phi|1 + (\vec{\sigma}\vec{p} - q\vec{A})\frac{1}{(2Mc)^2}(\vec{\sigma}\vec{p} - q\vec{A})|\phi\rangle$$

In the last term the magic Eq. 14.5 has been used.

14.5 The non-relativistic limit: Pauli equation

Let us now form the non-relativistic limit by setting $c = \infty$ in Eq. 14.6. In that case the relativistic "mass" M is equal to the rest mass m_0, which is constant in space. Thus we obtain the so-called **Pauli equation**.

PAULI EQUATION

$$\Big[\frac{(\vec{p} - q\vec{A})^2}{2m_0} + q\Phi - \frac{q}{m_0}\vec{S}\vec{B} - \epsilon\Big]|\phi\rangle = 0 \tag{14.7}$$

We have used here that

$$(\vec{p} - q\vec{A}) \times (\vec{p} - q\vec{A}) = -q(\vec{p} \times \vec{A} + \vec{A} \times \vec{p}) = i\hbar q(\vec{\nabla} \times \vec{A} + \vec{A} \times \vec{\nabla}) = i\hbar q\vec{B}$$

$\vec{B} = \vec{\nabla} \times \vec{A} + \vec{A} \times \nabla$ is the magnetic field. Here we have expressed the magnetic field as operator equation, where ∇ not only acts on $\vec{A}$ but also any function tow which this operator is applied to. The magnetic field obtained this way is a simple vector function of real space.

The spin is written as

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices.

Equation 14.7 differs from the Schrödinger equation only by the **Stern-Gerlach term** in the Hamiltonian[3]

$$\hat{H}_{SG} = -\frac{q}{m}\hat{\vec{S}}\vec{B}(\hat{\vec{r}})$$

which describes the energy of a magnetic moment $\vec{m} = \frac{q}{m}\vec{S}$ in a magnetic field $\vec{B}$. Because the spin of an electron has the eigenvalues $\pm\frac{\hbar}{2}$, the magnetic moment of the electron is $\mu_b = \frac{\hbar q}{2m_0}$. This fundamental constant is called the **Bohr magneton**. The magnetic moment in the Dirac equation is the origin of interpreting the spin as the rotation of a particle about its own axis, which is a strange concept for a point particle.

For a classical particle rotating about a center, magnetic moment and angular momentum are related by the relation $\vec{m} = \frac{q}{2m}\vec{L}$. The ratio between magnetic moment and angular momentum is called the **gyromagnetic moment**. For a classical particle the gyromagnetic moment is therefore $\gamma = \frac{q}{2m}$. For other charge distributions the ratio between angular momentum and magnetic moment differs, which is taken into account by the **g-factor** g so that in general $\vec{m} = g\frac{q}{m}\vec{L}$. Comparison with the Pauli equation indicates that the g-factor of the electron is $g_e = 2$ and the gyromagnetic ratio of the electron is $\gamma_e = g_e\frac{q}{2m} = \frac{q}{m}$. The magnetic moment of the electron is $g_e = 2.0023$. The deviation from the value 2 is explained by quantum electrodynamics.

Consider a classical particle circling about the origin in a constant magnetic field $\vec{B}$. The vector

[3] It is disturbing that the Pauli equation is a two-component equation, while the Schrödinger equation only has a single component. In most cases, one can decouple the two components of the Pauli equation by choosing the magnetic field in the z-direction. In that case there are no additional terms that couple spin-up and spin-down components. Thus one arrives at a description that distinguishes between spin-up and spin-down electrons. Each fulfills a Schrödinger equation, albeit with different potentials.

potential for a constant magnetic field can be written as $\vec{A} = \frac{1}{2}\vec{B} \times \vec{r}$. The energy can be written as

$$\begin{aligned}
E &= \frac{1}{2m}(\vec{p} - q\vec{A})^2 = \frac{1}{2m}(\vec{p} - \frac{q}{2}\vec{B} \times \vec{r})^2 \\
&= \frac{p^2}{2m} - \frac{q}{2m}\vec{p}(\vec{B} \times \vec{r}) + \frac{q^2}{8m}(\vec{B} \times \vec{r})^2 \\
&= \frac{p^2}{2m} - \frac{q}{2m}\vec{B}(\vec{r} \times \vec{p}) + \frac{q^2}{8m}(\vec{B} \times \vec{r})^2 \\
E(\vec{B}) &= E_0 - \frac{q}{2m}\vec{L}\vec{B} + O(B^2)
\end{aligned}$$

Now consider the limit where the orbit of the particle goes to zero, so that we obtain a point particle, and the angular momentum approaches the spin of the electron, that is $|L| \to \hbar/2$. We obtain to first order $E(B) = E_0 - \frac{q\hbar}{4m}\vec{B}$, assuming that the magnetic moment is parallel to the magnetic field. Thus we obtain half the value of that obtained from the Dirac equation. The factor missing is called the g-factor, and is considered an anomaly. We can attribute it to the fact that the classical particle in our model becomes itself relativistic as the radius is shrunk to zero.

If the magnetic field is parallel to the z-axis, the spin-up and spin-down components are independent. Each component is a solution of a Schrödinger equation, with the only difference that spin-up and spin-down components experience different potentials $V_\uparrow = q\Phi - \mu_B B_z$ and $V_\uparrow = q\Phi + \mu_B B_z$. If the electron is free to occupy the state with lowest energy, the spin will align parallel and not antiparallel to the magnetic field.

Just as we have determined the Pauli equation for the electrons from the Dirac equation for energies near $E = +m_0c^2$ after forming the non-relativistic limit, we can similarly determine a Pauli equation for the positrons, that is the solutions with energies near $E = -m_0c^2$. This is done explicitly in appendix M.2 on p. 332. There we also show that the motion of a positron corresponds indeed to a particle with the opposite charge of the electron.

14.6 Relativistic corrections

Relativistic effects are quite large near the nucleus, where the kinetic energy of an electron approaches infinity, and the velocity of the electrons approaches that of light.

Relativistic corrections can be observed macroscopically in every compass. A ferromagnet could not produce permanent magnetization direction without spin-orbit coupling. Spin-orbit coupling links the magnetization to the crystal geometry. In a compass without spin orbit coupling, the magnetization direction inside the material would orient towards the north-pole, but the needle would stay put.

In order to separate out the relativistic effects, let us introduce a new dimension-less variable

$$D = \frac{m}{M} - 1 \Rightarrow \frac{1}{2M} = \frac{1}{2m_0} + \frac{D}{2m_0}$$

D vanishes in the nonrelativistic limit $c \to \infty$, where $M = m$.

By replacing $\frac{1}{2M}$ by $\frac{1}{2m_0} + \frac{D}{2m_0}$ in the two-component Dirac equation Eq. 14.6, and after collecting all terms proportional to D, we obtain the relativistic corrections W to the Pauli equation

$$W = \frac{1}{2m_0}\Big\{(\vec{p} - q\vec{A})D(\vec{p} - q\vec{A}) + i\vec{\sigma}\Big((\vec{p} - q\vec{A})D \times (\vec{p} - q\vec{A})\Big)\Big\} \tag{14.8}$$

The two component Dirac equation can then be written as $(\hat{H}_0 + W - \epsilon)|\phi\rangle = 0$, where $\hat{H}_0$ is the Pauli Hamiltonian and W are the relativistic corrections. Note, that D contains not only a spatial dependence, but also an energy dependence. If we only consider the lowest order of the energy dependence, we can remove the explicit energy dependence.

Before we continue, let us investigate some properties of the function D

- As mentioned before, D vanishes in the non-relativistic limit.
- close to the relativistic limit that is for sufficiently slow particles, D has the form

$$D = -\frac{\epsilon - V}{2m_0c^2} + O(\frac{1}{c^4}) = -\frac{p^2}{4m_0^2c^2} + O(\frac{1}{c^4}). \tag{14.9}$$

  where $V = q\Phi$. The function D is negative and measures approximately the ratio of kinetic energy and the rest-energy of the particle.
- Let us consider the spatial dependence of D for an electron near a nucleus: We use $V(\vec r) = q\Phi = -Ze^2/(4\pi\epsilon_0|r|)$ and $\epsilon = 0$ as model for the potential in the center of an atom.

$$\begin{aligned} D &= \frac{m_0}{M} - 1 = \frac{1}{1 + \frac{\epsilon - V}{2m_0c^2}} - 1 = -\Big(1 + \frac{2m_0c^2}{\epsilon - V}\Big)^{-1} \\ &\approx -\Big(1 - \frac{8\pi\epsilon_0 m_0 c^2|r|}{Ze^2}\Big)^{-1} \approx -1 + 2\frac{|\vec r|}{r_T} + O(r^2) \end{aligned}$$

  where

$$r_T = \frac{Ze^2}{4\pi\epsilon_0 m_0 c^2}$$

  is the so-called Thomson radius. The Thomson radius gives an indication about the size of the region near the nucleus, where relativistic effects are important. The Thomson radius lies in the range of $10^{-4} - 10^{-2}$ a_0, where $a_0 \approx 0.529 \times 10^{-10}$ m is the Bohr radius. As a comparison, a typical atomic radius is in the range of 3 a_0.

Let us now return to the relativistic corrections of Eq. 14.8. To simplify the discussion we will introduce a number of approximations. As the first simplification we ignore the dependence of the relativistic corrections on the magnetic field. The idea is that usually both, relativistic corrections and the magnetic field, can be considered as small perturbations, so that they can be treated independently to first order.

We obtain

$$W = \frac{1}{2m_0}\Big\{\vec pD\vec p + i\vec\sigma\Big(\vec pD \times \vec p\Big)\Big\} \tag{14.10}$$

The last term in Eq. 14.10 will be identified with the spin-orbit coupling. The first term can be divided into two terms that will be identified with the Darwin term and the mass-velocity correction. We rewrite $\vec pD\vec p$ in two different ways and average the resulting expressions.

$$\begin{aligned} \vec pD\vec p &= \vec p[D, \vec p] + \vec p^2D = -\vec p[\vec p, D] + \vec p^2D = -[\vec p, [\vec p, D]] - [\vec p, D], \vec p + \vec p^2D \\ \vec pD\vec p &= [\vec p, D]\vec p + D\vec p^2 \\ \vec pD\vec p &= -\frac{1}{2}([\vec p, [\vec p, D]] + \frac{1}{2}(\vec p^2D + D\vec p^2) \\ &= \frac{\hbar^2}{2}(\nabla^2D) + \frac{1}{2}(\vec p^2D + D\vec p^2) \end{aligned}$$

where $(\nabla^2D(\vec r))$ is considered as a function and not an operator. [4] p^2 on the other hand is considered in both cases as an operator. We can rewrite the relativistic corrections as

$$\begin{aligned} W &= \frac{\hbar^2}{4m_0}(\vec\nabla^2D) + \frac{1}{4m_0}(\vec p^2D + D\vec p^2) + \frac{i}{2m_0}\vec\sigma\Big(\vec{\vec p}D \times \vec{\vec p}\Big) && (14.11) \\ &= W_D + W_{mv} + W_{LS} && (14.12) \end{aligned}$$

[4]Unfortunately the notation of operator equations is not completely consistent. Often, an additional pair of parentheses is used to indicate that the operator, in this case ∇^2 acts only on the function(s) inside the same pair of parentheses and not further onto any function in the operator or standing to the right of the operator. The true meaning must be obtained in the end from the context of the equation. It is always possible, even though in some cases utterly inconvenient to rewrite an operator equation such that every operator acts on everything standing to the right of it.

where the first term W_D is the Darwin term, W_{mv} is the mass-velocity correction and W_{LS} is the spin-orbit coupling. We will investigate the three contributions separately.

14.6.1 Spin-Orbit coupling

The last term of the relativistic corrections in Eq. 14.12 is the so-called **spin-orbit coupling**

$$W_{LS} = \frac{1}{2m_0} i\vec{\sigma}\Big(\vec{p}D \times \vec{p}\Big) \tag{14.13}$$

We will now introduce a some assumptions to allow a physical interpretation.

Let us assume that the electrostatic potential is spherically symmetric. As a result, also D is spherically symmetric, so that we can write

$$\vec{p}D = [\vec{p}, D]_- + D\vec{p} = \frac{\hbar}{i}(\partial_r D)\frac{\vec{r}}{|r|} + D\vec{p}$$

where we consider $(\partial_r D)$ as a single function. Note again the inconsistency in the notation, because one might expect that the radial derivative would need to be applied to any function standing to the right of D as well. Here and in the following we consider $(\partial_r D)$ as a symbol for a function.

Thus we find

$$\begin{aligned} W_{LS} &= \frac{1}{2m_0} i\vec{\sigma}\Big(\frac{\hbar}{i}\frac{(\partial_r D)}{|r|}\vec{r}\times\vec{p} + D\vec{p}\times\vec{p}\Big) = \frac{1}{2m_0} i\vec{\sigma}\frac{\hbar}{i}\frac{(\partial_r D)}{|r|}\vec{r}\times\vec{p} \\ &= \frac{(\partial_r D)}{m_0|r|}\vec{L}\vec{S} \end{aligned}$$

where we used that $\hbar\vec{\sigma}/2 = \vec{S}$ is the spin operator and $\vec{r}\times\vec{p} = \vec{L}$ is the angular momentum operator.

The energy splitting resulting from spin-orbit coupling in first order perturbation theory is

$$\Delta\epsilon_{LS} = \langle\psi|\frac{1}{m_0}\frac{(\partial_r D)}{|r|}\vec{L}\vec{S}|\psi_n\rangle = \pm\frac{\hbar^2\ell}{2m_0}\langle\psi|\frac{(\partial_r D)}{|r|}|\psi_n\rangle$$

The spin-orbit coupling splits an atomic energy level with an angular momentum quantum number ℓ into a upper level, where the spin is parallel to the orbital angular momentum, and a lower level with antiparallel spin and orbital momentum. Because the total angular momentum for the parallel spin level is $j = \ell + \frac{1}{2}$ the degeneracy of this upper orbital is $2j+1 = 2\ell+2$. The degeneracy of the lower level is 2ℓ.

For historical reasons one often finds the expression for the spin-orbit coupling in a different form. Here one employs first the assumption that the relativistic effects are small, that is $D = -\frac{\epsilon-V}{2m_0c^2} + O(\frac{1}{c^4})$, then one introduces the potential of the hydrogen atom and the assumption that the energy of the electron is small, that is $\epsilon = 0$. This yields

$$\Delta\epsilon_{LS} \approx \frac{Ze^2\hbar^2\ell}{16\pi\epsilon_0 m_0^2c^2}\langle\psi|\frac{1}{r^3}|\psi\rangle$$

This is a good (or bad) example how one can make incompatible approximations and still get away: Without corrections, D is not singular. However, we now used the approximation of small relativistic effects, that is relatively slow particles in combination with infinite potential energy, that is fairly fast particles, which lead us to a divergent expression for D near the nucleus. This is clearly nonsense! What saves the construction, is that (1) the $\ell = 0$ term vanishes, and (2) for other angular momenta the wave functions vanish sufficiently rapid near the nucleus, that they compensate the divergent term in the integral for the energy splitting.

The spin-orbit splitting is important for the conducting properties of semiconductors with heavier elements such as GaAs. Here the top of the valence band, which has p character is split into 4 bands with spin and orbital angular momentum parallel and two bands, where they are antiparallel. The spin-orbit splitting of these levels in silicon is 0.044 eV and in the heavier Ge it is already 0.3 eV.

The spin-orbit term is also visible in core level spectroscopy, where the core levels are split due to spin-orbit coupling into two multiplets that are separated by several eV.

14.6.2 Darwin term

In order to make the Darwin term[47]

$$W_D = \frac{\hbar^2}{4m_0}(\vec{\nabla}^2 D) \tag{14.14}$$

more transparent, let us now introduce again our usual approximations. First we keep only the leading order in $\frac{1}{c}$, so that $D = \frac{V-\epsilon}{2m_0c^2}$. Note that this approximation is not valid near the nucleus. I list it here, because that is what has been done historically. We obtain

$$W_D = \frac{\hbar^2 e(\vec{\nabla}^2\Phi)}{8m_0^2c^2} + O(\frac{1}{c^4})$$

We can simplify this expression further using the Poisson equation $\vec{\nabla}^2\Phi = \frac{\rho}{\epsilon_0}$.

If we specialize our result further to the Coulomb potential of the nucleus $V = q\Phi = -\frac{Ze^2}{4\pi\epsilon_0|r|}$ and $\epsilon = 0$, we obtain

$$W_D \approx \frac{\hbar^2 Z^2 e^2}{8\epsilon_0 m_0^2 c^2}\delta(r)$$

Thus the Darwin term acts only on s-wave function and shifts those up in energy.

14.6.3 Mass-velocity term

In order to give a physical interpretation to the mass-velocity

$$W_{mv} = \frac{1}{4m_0}(\vec{p}^2 D + D\vec{p}^2) \tag{14.15}$$

from Eq. 14.12, we use

$$\begin{aligned}
\epsilon - V|\psi\rangle &\approx \frac{\vec{p}^2}{2m_0}|\psi\rangle \\
\Rightarrow \langle\psi|D|\psi\rangle &\approx \langle\psi| - \frac{\vec{p}^2}{4m_0^2c^2}|\psi\rangle \\
\Rightarrow \langle\psi|W_{MV}|\psi\rangle &\approx \langle\psi| - \frac{\vec{p}^4}{8m_0^3c^2}|\psi\rangle
\end{aligned}$$

where we use the wave functions of the Pauli equation, implying that we only consider the leading order of the correction in $\frac{1}{c}$. The classical version of the mass velocity term is the leading term of the relativistic energy momentum relation.

$$\epsilon(p) = \sqrt{m_0^2c^4 + p^2c^2} - m_0c^2 = \frac{p^2}{2m_0} - \frac{p^4}{8m_0^3c^2} + O(\frac{1}{c^4})$$

where we used $\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + O(x^3)$

Chapter 15

Many Particles

Sofar, quantum theory was actually a theory of classical fields. Just as we describe light by electromagnetic fields, we introduced fields to describe other particles such as electrons. The new ingredients are the interpretation of the field amplitude as probability amplitude for particles to occur, and that we consequently normalized this probability to one.

15.1 Bosons

Let us return to the continuous linear chain that we investigated in one of the first sections. We showed there, that we could form a quantum theory by constructing wave packets that behave like classical particles. However, since the linear chain itself was build up from particles, we should treat every bead of the chain itself as a quantum particle. This idea is the central idea of **second quantization** and the treatment of many particles.

15.1.1 Quantum mechanics as classical wave theory

Let us consider the Lagrangian of the hanging linear chain, which can be also used to describe a relativistic particle. There are several reasons for choosing this model: we can look at it both as a classical model and a quantum mechanical problem. (1) If we remove the "mass" term that results from hanging the linear chain, we have a model for a phonon in a crystal. (2) If we ignore the fact that the displacements are real, the displacement field follows the same equations as the Klein-Gordon field, which describes relativistic spin-0 particles, such as mesons. Furthermore we avoid the intricacies of complex field amplitudes in the derivations.

We start out from the Lagrangian Eq. 3.31, which has been derived early in this booklet.

$$\mathcal{L}=K\int_0^L dx\Big(\frac{1}{c^2}(\partial_t\phi)^2-(\partial_x\phi)^2-\frac{m_0^2c^2}{\hbar^2}\phi(x,t)^2\Big)$$

The first step is to decompose $\phi(x,t)$ into eigen modes.[1] First we derive the Euler-Lagrange equations, which are the classical field equations such as the Schrödinger equation. Here we obtain:

$$\Big(-\frac{1}{c^2}\partial_t^2+\partial_x^2-\frac{m_0^2c^2}{\hbar^2}\Big)\phi(x,t)=0$$

We obtain the time independent equation using the ansatz $\phi(x,t)=u_j(x)\mathrm{e}^{-i\omega_j t}$

$$\Big(\frac{\omega_j^2}{c^2}+\partial_x^2-\frac{m_0^2c^2}{\hbar^2}\Big)u_j(x)=0$$

[1]We could also have staid in the real space formulation, but that would have made the derivation a bit more complex. Here I rather choose to transform first into Eigen modes and later transform back into real space.

The solutions are

$$u_j(x) = \frac{1}{\sqrt{L}}\sin(k_j x) \quad \text{with} \quad k_j = \frac{2\pi}{L}j$$
$$\hbar\omega_j = \sqrt{m_0^2c^4 + (\hbar k_j c)^2}$$

where we use periodic boundary conditions $\phi(x) = \phi(x + LL)$ and the ortho-normalization condition $\langle u_j|u_{j'}\rangle = \int_0^L dx u_j(x)u_{j'} = \delta_{j,j'}$.

There is one tricky point: We only keep the solution $\omega > 0$ even though the equation has a positive and a negative branch. This is analogous to the harmonic oscillator where the classical motion is built from solutions $x(t) = A\mathrm{e}^{+i\omega_0 t} + B\mathrm{e}^{-i\omega_0 t}$. Nevertheless, we only consider positive frequency for the harmonic oscillator.

Now we need to introduce the mode decomposition $\phi(x,t) = \sum_j u_j(x)\phi_j(t)$ into the Lagrangian, which we rewrite with the help of Gauss theorem and the Euler-Lagrange equations obtained above.

$$\begin{aligned}
\mathcal{L} &= K\int_0^L dx\Big(\frac{1}{c^2}(\partial_t\phi)^2 + \phi(\partial_x^2 - \frac{m_0^2c^2}{\hbar^2})\phi\Big)\\
&= K\sum_{n,n'}\Big(\frac{1}{c^2}(\partial_t\phi_j)(\partial_t\phi_j')\int_0^L dx u_j(x)u_j'(x)\\
&\quad +\phi_j\phi_j'\int_0^L dx u_j(x)(\partial_x^2 - \frac{m_0^2c^2}{\hbar^2})u_j'(x)\Big)\\
&= \frac{K}{c^2}\sum_j\Big((\partial_t\phi_j)^2 - \omega_j^2\phi_j^2\Big)
\end{aligned}$$

In the next step we need to find the Hamilton function for the classical displacement field. We obtain the the canonical momentum of the displacement field amplitude $\phi_j(t)$,

$$\Pi_j = \frac{\partial\mathcal{L}}{\partial\dot{\phi}_j} = \frac{2K}{c^2}\partial_t\phi_j \Rightarrow \partial_t\phi_j = \frac{c^2}{2K}\Pi_j$$

and set up the Hamilton function in the usual way

$$\begin{aligned}
H(\Pi_j,\phi_j) &= \sum_j \Pi_j\partial_t\phi_j - \mathcal{L}(\phi_j\partial_t\phi_j)\\
&= \frac{c^2}{4K}\sum_j\Big(\Pi_j^2 + \frac{4K^2\omega_j^2}{c^4}\phi_j^2\Big) = E
\end{aligned}$$

15.1.2 Second quantization

Now we quantize the displacement field $\phi(x)$. The procedure is completely analogous to the first quantization, where we made the transition from a position coordinate x to a wave function $\phi(x)$ on the coordinate space. Here we make the transition from a displacement field or probability amplitude to a wave function Φ defined on the space spanned by the probability amplitudes. This implies that the wave function amplitude $\phi(x,t)$ does not have a well defined value but is itself described by a wave function $\Phi[\phi(x)]$. Note that Φ is a functional of the displacement field. For the linear chain, this is a natural step, because it corresponds to the first quantization of the beads. If however we interpret $\phi(x,t)$ as the probability amplitude of a quantized point particle, things become different.

We use the correspondence principle

$$E \to i\hbar\partial_t \qquad \Pi_j \to \frac{\hbar}{i}\partial_{\phi_j} \tag{15.1}$$

to convert the Hamilton function into a Hamilton operator.

We obtain a Schrödinger equation

$$i\hbar\partial_t|\Phi\rangle = \mathcal{H}|\Phi\rangle = \frac{c^2}{4K}\sum_j\Big(-\hbar^2(\partial_{\phi_j})^2 + \frac{4K^2\omega_j^2}{c^4}\phi_j^2\Big)|\Phi\rangle$$

where Φ is a wave function that depends on the displacements of all the eigen modes $\Phi(\phi_1, \phi_2, \ldots)$. Note here, that the difference between the Hamilton operator used here and the Hamilton operator $H(p, x) = -\frac{\hbar^2}{2m}\partial_x^2 + V(x)$, that defines the "classical field equations" such as the Schrödinger equation.

Let us try to give some meaning to this Schrödinger equation. What we find is a harmonic oscillator for each vibrational mode. We know how to deal with this from Sec. 5.6. By comparing with Eq. 5.30, we identify the mass of the harmonic oscillator with $m = \frac{2K}{c^2}$ and the frequency with $\omega_0 = \omega_j$. We introduce the ladder operators a and $\hat{a}^\dagger$ analogous to Eq. 5.33.

$$\hat{a}_j = i\sqrt{\frac{c^2}{4\hbar K\omega_j}}\hat{\Pi}_j + \sqrt{\frac{K\omega_j}{\hbar c^2}}\hat{\phi}_j$$

$$\hat{a}_j^\dagger = -i\sqrt{\frac{c^2}{4\hbar K\omega_j}}\hat{\Pi}_j + \sqrt{\frac{K\omega_j}{\hbar c^2}}\hat{\phi}_j$$

Note that $\hat{\Pi}_j = \frac{\hbar}{i}\partial_{\phi_j}$ is an operator and no more a classical field with a well defined value at each position.

With the help of $\epsilon_j = \hbar\omega_j$, the Hamilton operator can be written as

$$\hat{\mathcal{H}} = \sum_j |u_j\rangle\epsilon_j\Big(\hat{a}_j^\dagger\hat{a}_j + \frac{1}{2}\Big)\langle u_j|$$

What we get are discrete energy levels with values

$$E(n_j) = \sum_j \epsilon_j(\frac{1}{2} + n_j).$$

The energy is discretized and comes in bunches, where each **quantum** carries the energy corresponding to the dispersion relation $\epsilon(k_j) = \hbar k_j$. Thus we can interpret each excitations of a vibrational mode as particle with a wave function with the corresponding eigen mode of the hanging linear chain. We obtained the particle properties, namely discretized energy quanta from a pure field theory!

All states can be built up from a **vacuum state** $|\mathcal{O}\rangle$, which corresponds to no particles at all. Note that the vacuum state is not the zero state! A state with n_1 particles in the first vibrational mode, n_2 particles in the second and so on, can be constructed as

$$|n_1, n_2, \ldots\rangle = \frac{\prod_j (a_j^\dagger)^{n_j}}{\prod_j n_j!}|\mathcal{O}\rangle$$

We have actually used here two equivalent notations for a given state, the **number representation** on the left side and the one using creation operators. We have extended Hilbert space used for one-particle states, and formed the so-called **Fock space**, which contains the Hilbert space for one particle, the one for two particles and so on.

Now we can understand the double-slit experiment for photons. For the quantized Maxwell's equation the double-slit experiment has the following interpretation: In order to find out if electromagnetic radiation has been absorbed, we need to measure the intensity of the electromagnetic field before and after absorption. The intensity is related to the particle number operator $\sum_j a_j^\dagger a_j$. The resulting eigenvalues, however only change by integer amounts. If the intensity absorbed is very large, the changes of the intensity relative to the total intensity are so small that they become irrelevant, so that we obtain the classical interpretation of the intensity. **Editorial remark: Where are the anti-particles?**

15.1.3 Coordinate representation

Editorial remark: This is pretty strong!

We can also transform this into the coordinate representation using

$$\phi(x,t) = \sum_j u_j(x)\phi_j(t)$$
$$\Pi(x,t) = \sum_j u_j(x)\Pi_j(t)$$

Let us consider the commutator relations among the two operator fields. We use

$$[\Pi_j, \phi_{j'}] = \frac{\hbar}{i}\delta_{j,j'} \qquad [\phi_j, \phi_{j'}] = 0 \qquad [\Pi_j, \Pi_{j'}] = 0$$

which is a direct consequence of the correspondence principle.

We obtain

$$\begin{aligned}
[\Pi(x), \phi(x')] &= \sum_{j,j'} u_j(x)u_{j'}(x')[\Pi_j, \phi_{j'}] \\
&= \sum_j u_j(x)u_j(x') = \delta(x-x')
\end{aligned}$$

and therefore

$$[\Pi(x), \phi(x')] = \frac{\hbar}{i}\delta(x-x') \qquad [\phi(x), \phi(x')] = 0 \qquad [\Pi(x), \Pi(x')] = 0$$

The creation and annihilation operators can be transformed between the mode representation and the real space transformation.

$$a^\dagger(x) = \sum_j u_j^*(x)a_j^\dagger$$
$$a(x) = \sum_j u_j(x)a_j$$

The inverse transformation is

$$a_j^\dagger = \int dx u_j(x)a^\dagger(x)$$
$$a_j = \int dx u_j^*(x)a(x)$$

Hence the commutator relations of the creation and annihilation operators are

$$[a(x), a^\dagger(x')] = \delta(x-x') \qquad [a^\dagger(x), a^\dagger(x')] = 0 \qquad [a(x), a(x')] = 0$$

Let us now consider the Hamilton operator and transform it to real space

$$\begin{aligned}
H &= \sum_j \epsilon_j(a_j^\dagger a_j + \frac{1}{2}) \\
&= \sum_j \frac{\epsilon_j}{2}\Big((\int dx u_j(x)a^\dagger(x))(\int dx' u_j^*(x')a(x')) + (\int dx u_j^*(x)a(x))(\int dx' u_j(x')a^\dagger(x'))\Big) \\
&= \int dx \int dx' a^\dagger(x)\Big(\sum_j u_j(x)\epsilon_j u_j^*(x')\Big)a(x) + \int dx \int dx' a(x)\Big(\sum_j u_j^*(x)\epsilon_j u_j(x')\Big)a^\dagger(x) \\
&= \frac{1}{2}\int dx\Big[a^\dagger(x)\sqrt{m_0^2c^4 + c^2\partial_x^2}a(x) + a(x)\sqrt{m_0^2c^4 + c^2\partial_x^2}a^\dagger(x)
\end{aligned}$$

Here we used

$$
\begin{aligned}
\int dx \int dx' f(x) u_j^*(x) \epsilon_j u_j(x') g(x') &= \sum_j \langle f|u_j\rangle\langle u_j|H|u_j\rangle\langle u_j|g\rangle \\
&= \langle f|\Big(|u_j\rangle\langle u_j|\Big) H\Big(\sum_j |u_j\rangle\langle u_j|\Big)|g\rangle \\
&= \langle f|H|g\rangle
\end{aligned}
$$

Here a space derivative appears under a square root, which is dealt with in the usual way, namely as power series expansion of the square root.

Note, that we deal now with wave functions on an infinite dimensional space. Thus it is practically impossible to write down such a wave function. If we resort to the abstract operator algebra the expressions are quite intuitive.

Now some serious business: let us consider free space, then we are left with the zero point motion of the particle coordinate, which carries an infinite energy

$$
E = \int \frac{dk}{2\pi} \frac{1}{2} \epsilon(k)
$$

Since $\epsilon(k)$ grows with increasing k the vacuum has infinite energy. Should we worry? Yes! We know from the theory of relativity that energy is related to mass, and mass results in gravitational forces. They should be observable.

This turned out to be one of the major puzzles in physics, namely that the two main theories, general theory of relativity, which describes gravitational forces, and quantum theory seem to be incompatible.

15.2 Fermions

Fermions are constructed similarly. The main difference is that instead of introducing the ladder operators of the harmonic oscillator for each vibrational mode we introduce those of the two state system. Again the Fermionic waves come in bunches, but only one Fermion can be created in one vibrational mode of the linear chain.

Unfortunately, I do not know of a good classical analogue for Fermions and people say there is none. Therefore, you need to rely on your imagination.

Fermions could solve the problem of infinite zero point energy of the bosons. As Fermions have negative zero-point energy, the two energies would cancel if they would exist in equal numbers. This is the basic idea of super-symmetry, which assumes that for every Boson, there is also a Fermion, but now we have ventured very far from the material of an elementary course of quantum mechanics.

15.3 Final remarks

We only scratched the surface. We have seen that if we use the correspondence principle for the wavefunction amplitudes we obtain discretized excitations, which we interpret as particles. The same principle, which made waves out of point particle, made particles out of continuous wave functions. Thus we can now describe all the puzzling observations in the double-slit experiment both for electrons and for photons. The price is, however, a theory of some complexity: In the end we deal with wave functions on infinite dimensional spaces!

For further reading on quantum field theory, I recommend the book by Haken [7]. It should be noted that quantum field theory, the description of many particle states, comes in many different appearances, such as Green's functions, Feynman diagrams, creation and annihilation operators etc., which all have their own virtues and disadvantages.

Chapter 16

Exercises

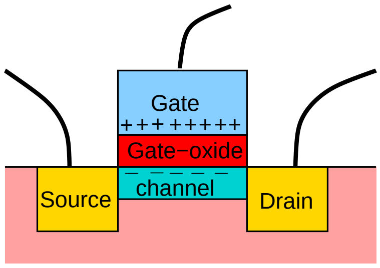
16.1 Interference

Describe the mechanism of interference

16.2 Penetration of electrons into an insulator

Editor: This is only a sketch

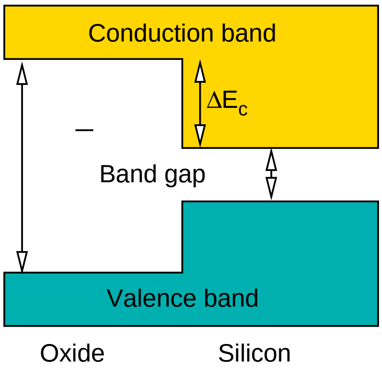
Most transistors in today's computers are so-called metal-oxide-semiconductor Field-effect transistors (MOSFET). A transistor is a switch, where a current between source and drain



One important problem is that there are defects in the gate oxide, which have electronic states that can be filled by electrons. Electrons can tunnel from the channel into these defect levels. These correspond to another double layer of charge. This affects the capacity of the capacitor and with it the characteristics of the transistor.

We want to estimate how deep the wave functions penetrate from the semiconductor into the oxide, because from the wave function amplitude we can estimate how fast the defects will charge up as function of the distance from the semiconductor.

We describe the conduction band states of the semiconductor as a free electrons with an effective mass of $m^*_{Si} = 0.5m_e$, where m_e is the mass of an electron in vacuum. The oxide we also describe as free electron states with an effective mass of $m^*_{SiO_2} =??m_e$. The conduction band of the oxide lies about 3 eV above that of the semiconductor. A defect level is located at a distance of 0.2 nm from the silicon channel and is located 0.2 eV above the conduction band of the semiconductor.



Determine the probability density of a wave function from silicon at the defect level.

16.3 Quantized conductance

Editor: This is only a sketch

16.4 Translation operator from canonical momentum

Problem:

Use the relation $\hat{S} = \mathrm{e}^{-\frac{i}{\hbar}\hat{p}x}$ with the momentum

$$\hat{p} = \int dx\, |x\rangle \frac{\hbar}{i}\partial_x \langle x| \tag{16.1}$$

to show that $\hat{S}\Psi(x) = \Psi(x-\Delta)$

Solution:

The shift operator given above is a function of the momentum operator, which is defined by its power series expansion. Whenever the function depends only on one operator one best changes the representation into eigenstates of that operator. Thus we need to transform the momentum operator into the basis of its eigenstates. Most of the derivation consists of introducing again the momentum representation, which has been worked out previously.

The eigenstates of the momentum operator are denoted by $|p\rangle$. They fulfill the eigenvalue equation

$$\hat{p}|p\rangle = |p\rangle p \tag{16.2}$$

and are normalized as follows:

$$\langle p|p'\rangle = 2\pi\hbar\delta(p-p') \tag{16.3}$$

The inclusion of the factor $2\pi\hbar$ is the convention used in this text. Other books often drop the factor.

The real-space wave functions corresponding to momentum eigenstates have the form

$$\langle x|p\rangle = \mathrm{e}^{\frac{i}{\hbar}px} \tag{16.4}$$

Eq. 16.4 is shown as follows: The eigenvalue equation Eq. 16.2 can easily be verified

$$\langle x|\hat{p}|p\rangle \stackrel{\text{Eq. 16.1}}{=} \int dx' \underbrace{\langle x|x'\rangle}_{\delta(x-x')} \frac{\hbar}{i}\partial_{x'} \underbrace{\langle x'|p\rangle}_{\mathrm{e}^{ipx'/\hbar}} = \int dx'\delta(x-x')\, p\mathrm{e}^{\frac{i}{\hbar}px'} = \mathrm{e}^{\frac{i}{\hbar}px}p = \langle x|p\rangle p$$

Now we need to check the normalization condition Eq. 16.3

$$\langle p|p'\rangle = \langle p|\overbrace{\int dx\ |x\rangle\langle x|}^{\hat{1}}|p'\rangle = \int dx\ \langle p|x\rangle\langle x|p'\rangle = \int dx\ \mathrm{e}^{\frac{i}{\hbar}(p-p')x}$$

Clearly the integral does not converge. Therefore, we introduce a limiting factor to the integrand

$$\begin{aligned}
\langle p|p'\rangle &= \lim_{\lambda\to 0}\left(\int_{-\infty}^{0} dx\ \mathrm{e}^{(\frac{i}{\hbar}(p-p')+\lambda)x} + \int_{0}^{\infty} dx\ \mathrm{e}^{(\frac{i}{\hbar}(p-p')-\lambda)x}\right) \\
&= \lim_{\lambda\to 0}\left(\left(\frac{i}{\hbar}(p-p')+\lambda\right)^{-1} - \left(\frac{i}{\hbar}(p-p')-\lambda\right)^{-1}\right) \\
&= \lim_{\lambda\to 0}\frac{-2\lambda}{\frac{-1}{\hbar^2}(p-p')^2-\lambda^2} \\
&= 2\hbar\lim_{\lambda\to 0}\frac{\hbar\lambda}{(p-p')^2+\hbar^2\lambda^2}
\end{aligned}$$

We recognize already that the function vanishes for $p \neq p'$ and that it diverges for $p = p'$ which suggests that the limes could be a delta function. What needs to be shown is that the integral over p is equal to a well-defined constant.

We use the integral formula taken from Bronstein[26]

$$\int dx \frac{1}{x^2+a^2} = \frac{1}{a}\mathrm{atan}(\frac{x}{a})$$

Thus we obtain

$$\begin{aligned}
\int_{-\infty}^{\infty} dp\langle p|p'\rangle &= 2\hbar\lim_{\lambda\to 0}\left[\mathrm{atan}(\frac{p-p'}{\hbar\lambda})\right]_{-\infty}^{\infty} = 2\hbar\left[\underbrace{\mathrm{atan}(\infty)}_{\pi/2} - \underbrace{\mathrm{atan}(-\infty)}_{-\pi/2}\right] \\
&= 2\pi\hbar
\end{aligned}$$

Thus, we can identify the scalar product as

$$\langle p|p'\rangle = 2\pi\hbar\delta(p-p')$$

which verifies the normalization condition Eq. 16.3.

The unity operator in momentum representation has the form

$$\hat{1} = \int \frac{dp}{2\pi\hbar}\ |p\rangle\langle p|$$

The form of the unity operator can be verified by applying it to an arbitrary state $|p\rangle$. Due to the superposition principle, it holds for any state if it holds for a complete set of states such as $|p\rangle$. The identity is shown as follows:

$$\hat{1}|p\rangle = \int \frac{dp'}{2\pi\hbar}|p'\rangle\langle p'|p\rangle = \int \frac{dp'}{2\pi\hbar}|p'\rangle 2\pi\hbar\delta(p-p') = |p\rangle$$

Now we can evaluate the shift operator in momentum representation

$$\hat{S}(\Delta) = \underbrace{\mathrm{e}^{-\frac{i}{\hbar}\hat{p}\Delta}}_{\hat{S}(\Delta)}\underbrace{\int \frac{dp}{2\pi\hbar}|p\rangle\langle p|}_{\hat{1}} = \int \frac{dp}{2\pi\hbar}\mathrm{e}^{-\frac{i}{\hbar}\hat{p}\Delta}|p\rangle\langle p| = \int \frac{dp}{2\pi\hbar}|p\rangle\mathrm{e}^{-\frac{i}{\hbar}p\Delta}\langle p|$$

Note that the exponential does not carry the momentum operator, but its eigenvalue, which is a number. Now we can determine the result of the shift operator acting on a wave function $\psi(x)$.

$$\begin{aligned}\hat{S}(\Delta)\psi(x) &= \langle x|\hat{S}(\Delta)|\psi\rangle = \langle x|\int \frac{dp}{2\pi\hbar}|p\rangle \mathrm{e}^{-\frac{i}{\hbar}p\Delta}\langle p||\psi\rangle \\ &= \int \frac{dp}{2\pi\hbar}\underbrace{\langle x|p\rangle}_{\mathrm{e}^{\frac{i}{\hbar}px}}\mathrm{e}^{-\frac{i}{\hbar}p\Delta}\underbrace{\langle p|\psi\rangle}_{\psi(p)} = \int \frac{dp}{2\pi\hbar}\,\mathrm{e}^{\frac{i}{\hbar}p(x-\Delta)}\psi(p)\end{aligned}$$

We compare the result with the original wave function expressed by its momentum representation

$$\begin{aligned}\hat{1}\psi(x) &= \langle x|\hat{1}|\psi\rangle = \langle x|\int \frac{dp}{2\pi\hbar}\,|p\rangle\langle p||\psi\rangle \\ &= \int \frac{dp}{2\pi\hbar}\underbrace{\langle x|p\rangle}_{\mathrm{e}^{\frac{i}{\hbar}px}}\underbrace{\langle p|\psi\rangle}_{\psi(p)} = \int \frac{dp}{2\pi\hbar}\,\mathrm{e}^{\frac{i}{\hbar}px}\psi(p)\end{aligned}$$

We observe directly that

$$\hat{S}\psi(x) = \psi(x-\Delta) \tag{16.5}$$

16.5 Rotator

Problem

Consider the motion of the two nuclei of mass m in a diatomic molecule with an equilibrium bond distance d and a force constant c.

1. Set up the Hamiltonian and divide it into a translational, a rotational and a vibrational part.
2. Determine explicitly the energy spectrum for the rotational motion.
3. Describe the complete spectrum of the molecule. and represent it in terms of dispersion relations.

16.6 Symmetrization of states

This exercise is a draft!

Consider three hydrogen atoms placed at the corners of an equilateral triangle. The wave function of one atom at the origin $\psi(\vec{r}) = \mathrm{e}^{-\lambda|\vec{r}|}$. Construct eigenstates of the rotation operator about the three-fold rotation axis, which are superpositions of the states of the three atoms. Only one wave function per atom is considered. Provide the coefficients of the three eigenstates and their eigenvalues.

16.7 Kronig-Penney model

Problem:

The Kronig Penney model[48] is probably the most simple model for the band structure of a solid.

The Dirac comb is a one-dimensional model system for a crystal of atoms. The potential of a single "Dirac atom" at position X_0 is a δ-function $V(x) = -V_0\delta(x - X_0)$.

Use the Bloch theorem to determine the eigenstates and the band structure of a particle in a the Dirac comb potential. The Dirac Comb potential is a periodic array of Dirac atoms arranged with a lattice constant of a.

Hint: The wave function in a δ-function like potential has a step in the derivative, which is obtained by inserting it in the one-dimensional Schrödinger equation

$$
\begin{aligned}
0 &= \lim_{\epsilon\to 0}\int_{-\varepsilon}^{\varepsilon} dx \left[-\frac{\hbar^2}{2m}\partial_x^2 - V_0\delta(x) - E\right]\psi(x) = \lim_{\epsilon\to 0}\left[-\frac{\hbar^2}{2m}\left(\partial_x|_\epsilon\,\psi - \partial_x|_{-\epsilon}\,\psi\right) - V_0\psi(0)\right] \\
\Rightarrow \quad & \left(\partial_x\psi\right)_{0^+} - \left(\partial_x\psi\right)_{0^-} + \frac{2mV_0}{\hbar^2}\psi(0) = 0
\end{aligned}
\tag{16.6}
$$

where x^+ stands for $x + \epsilon$ in the limit of vanishing positive ϵ.

Remark: This exercise is very instructive, but cannot completely be solved analytically. In the end use a graphical method to demonstrate the qualitative results.

Remark: Kronig and Penney use the opposite sign of V_0. This describes the atom as a box with a narrow barrier between atoms. Our model is that of a point-like atom with a wide barrier between atoms.

Solution

Potential

The potential is

$$
V(x) = -V_0 \sum_{n=-\infty}^{\infty} \delta(x - na)
$$

Bloch theorem

Now we can approach the crystal: We use Bloch theorem to express the wave function by a periodic part $u_k(x)$ and a phase factor e^{ikx}

$$
\psi_k(x) = u_k(x)\mathrm{e}^{ikx}
$$

The Schrödinger equation for the periodic part $u_k(x)$ is

$$
\begin{aligned}
0 &= \left[-\frac{\hbar^2}{2m}\partial_x^2 - V_0\sum_n \delta(x - na) - E\right] u_k(x)\mathrm{e}^{ikx} \\
&= \mathrm{e}^{ikx}\left[-\frac{\hbar^2}{2m}(\partial_x + ik)^2 - V_0\sum_n \delta(x - na) - E\right] u_k(x) = 0 \\
\Rightarrow \quad & \left[-\frac{\hbar^2}{2m}(\partial_x + ik)^2 - V_0\sum_n \delta(x - na) - E\right] u_k(x) = 0
\end{aligned}
$$

Set up conditions

Conditions for the wave function

- $u_k(x)$ is periodic, i.e $u_k(x) = u_k(x-a)$,
- $u_k(x)$ is continuous,
- $u_k(x)$ has a step in the first derivative at the position of the Dirac atom according to Eq. 16.6,
- $u_k(x)$ is a solution of the Schrödinger equation $[\frac{-\hbar^2}{2m}(\partial_x + ik)^2 - E]u_k(x) = 0$ in between the Dirac atoms, and
- $u_k(x)$ is normalized within the unit cell, i.e. $\int_0^a dx\, |\psi(x)|^2 = 1$.

Schrödinger equation for the periodic part

Let us start with the Schrödinger equation in between the Dirac atoms.

$$\left[\frac{-\hbar^2}{2m}(\partial_x + ik)^2 - E\right] u_k(x) = 0$$

Because THIS equation is translationally invariant we conclude that the solutions are a plane wave or exponential functions. This can also be concluded from the fact that it is a differential equation with constant coefficients.

Let us choose a partial solution $\exp(i\bar{G}x)$ and insert it into the Schrödinger equation above

$$0 = \left[\frac{-\hbar^2}{2m}(\partial_x + ik)^2 - E\right] \exp(i\bar{G}x) = \exp(i\bar{G}x)\left[\frac{-\hbar^2}{2m}(i\bar{G} + ik)^2 - E\right]$$

$$\Rightarrow \qquad E = \frac{\hbar^2}{2m}(k + \bar{G})^2 \qquad \text{or} \qquad \bar{G} = -k \pm \sqrt{\frac{2mE}{\hbar^2}} \tag{16.7}$$

We use the Ansatz for the wave function

$$u_k(x) = \left(A\mathrm{e}^{-i(k-G)x} + B\mathrm{e}^{-i(k+G)x}\right) \tag{16.8}$$

where we define G without a bar as

$$G \stackrel{\text{def}}{=} \sqrt{\frac{2mE}{\hbar^2}} \tag{16.9}$$

Continuity

The condition of continuity gives together with the condition of periodicity to

$$u(0^+) - u(0^-) = 0 \qquad \Rightarrow \qquad u(0^+) - u(a^-) = 0 \tag{16.10}$$

With the Ansatz from Eq. 16.8, we obtain conditions.

$$\begin{aligned} A\mathrm{e}^{-i(k-G)0^+} + B\mathrm{e}^{-i(k+G)0^+} - A\mathrm{e}^{-i(k-G)a^-} - B\mathrm{e}^{-i(k+G)a^-} &\stackrel{\text{Eq. 16.10}}{=} 0 \\ \Rightarrow \qquad \left(1 - \mathrm{e}^{-i(k-G)a}\right)A + \left(1 - \mathrm{e}^{-i(k+G)a}\right)B &= 0 \end{aligned}$$

$$\Rightarrow \qquad B = -\frac{1 - \mathrm{e}^{-i(k-G)a}}{1 - \mathrm{e}^{-i(k+G)a}}A \tag{16.11}$$

Thus the periodic part of our wave function has the form

$$u_k(x) = C\left[\left(1 - \mathrm{e}^{-i(k+G)a}\right)\mathrm{e}^{-i(k-G)x} - \left(1 - \mathrm{e}^{-i(k-G)a}\right)\mathrm{e}^{-i(k+G)x}\right) \tag{16.12}$$

where C is the normalization constant

Kink condition

Using periodicity, the kink-condition Eq. 16.6 leads to

$$
\begin{aligned}
&& \Big(\partial_x\psi\Big)_{0^+}-\Big(\partial_x\psi\Big)_{0^-}+\frac{2mV_0}{\hbar^2}\psi(0)=0\\
&\Rightarrow & \Big(\partial_x u_k(x)\mathrm{e}^{ikx}\Big)_{x=0^+}-\Big(\partial_x u_k(x)\mathrm{e}^{ikx}\psi\Big)_{0^-}+\frac{2mV_0}{\hbar^2}u_k(0)\mathrm{e}^{ik0}=0\\
&\overset{u_k(x-a)=u_k(x)}{\Rightarrow} & \mathrm{e}^{ik0^+}\Big((\partial_x+ik)u_k(x)\Big)_{x=0^+}-\mathrm{e}^{ik0^-}\Big((\partial_x+ik)u_k(x)\Big)_{a^-}+\frac{2mV_0}{\hbar^2}u_k(0)\mathrm{e}^{ik0}=0\\
&\Rightarrow & \Big((\partial_x+ik)u_k\Big)_{0^+}-\Big((\partial_x+ik)u_k\Big)_{a^-}+\frac{2mV_0}{\hbar^2}u_k(0)=0\\
&\overset{u_k(0)=u_k(a)}{\Rightarrow} & \Big(\partial_x u_k\Big)_{0^+}-\Big(\partial_x u_k\Big)_{a^-}+\frac{2mV_0}{\hbar^2}u_k(0)=0
\end{aligned}
$$

Now we insert the Ansatz Eq. 16.12, and drop the global factor C, which gives us an equation for G

$$
\begin{aligned}
0=&\Big[-i(k-G)\Big(1-\mathrm{e}^{-i(k+G)a}\Big)\mathrm{e}^{-i(k-G)0^+}+i(k+G)\Big(1-\mathrm{e}^{-i(k-G)a}\Big)\mathrm{e}^{-i(k+G)0^+}\Big]\\
&-\Big[-i(k-G)\Big(1-\mathrm{e}^{-i(k+G)a}\Big)\mathrm{e}^{-i(k-G)a^-}+i(k+G)\Big(1-\mathrm{e}^{-i(k-G)a}\Big)\mathrm{e}^{-i(k+G)a^-}\Big]\\
&+\frac{2mV_0}{\hbar^2}\Big[\Big(1-\mathrm{e}^{-i(k+G)a}\Big)\mathrm{e}^{-i(k-G)0}-\Big(1-\mathrm{e}^{-i(k-G)a}\Big)\mathrm{e}^{-i(k+G)0}\Big]\\
=&\Big[-i(k-G)\Big(1-\mathrm{e}^{-i(k+G)a}\Big)+i(k+G)\Big(1-\mathrm{e}^{-i(k-G)a}\Big)\Big]\\
&-\Big[-i(k-G)\Big(1-\mathrm{e}^{-i(k+G)a}\Big)\mathrm{e}^{-i(k-G)a^-}+i(k+G)\Big(1-\mathrm{e}^{-i(k-G)a}\Big)\mathrm{e}^{-i(k+G)a^-}\Big]\\
&+\frac{2mV_0}{\hbar^2}\Big[1-\mathrm{e}^{-i(k+G)a}-1+\mathrm{e}^{-i(k-G)a}\Big]\\
=&\Big[-i(k-G)\Big(1-\mathrm{e}^{-i(k+G)a}\Big)\Big(1-\mathrm{e}^{-i(k-G)a}\Big)+i(k+G)\Big(1-\mathrm{e}^{-i(k-G)a}\Big)\Big(1-\mathrm{e}^{-i(k+G)a}\Big)\\
&+\frac{2mV_0}{\hbar^2}\Big[-\mathrm{e}^{-i(k+G)a}+\mathrm{e}^{-i(k-G)a}\Big]\\
=&\Big[4i(k-G)\mathrm{e}^{-ika}\sin\Big(\frac{(k+G)a}{2}\Big)\sin\Big(\frac{(k-G)a}{2}\Big)-4i(k+G)\mathrm{e}^{-ika}\sin\Big(\frac{(k-G)a}{2}\Big)\sin\Big(\frac{(k+G)a}{2}\Big)\\
&+\frac{2mV_0}{\hbar^2}\Big[-2i\mathrm{e}^{-ika}\sin(Ga)\Big]\\
=&4i\mathrm{e}^{-ika}\Big[-2G\sin\Big(\frac{(k+G)a}{2}\Big)\sin\Big(\frac{(k-G)a}{2}\Big)-\frac{mV_0}{\hbar^2}\sin(Ga)\Big]
\end{aligned}
\tag{16.13}
$$

We use the identity

$$
\begin{aligned}
&\cos(x+y)=\mathrm{Re}[\mathrm{e}^{i(x+y)}]=\mathrm{Re}[\mathrm{e}^{ix}\mathrm{e}^{iy}]=\mathrm{Re}[\mathrm{e}^{ix}]\mathrm{Re}[\mathrm{e}^{iy}]-\mathrm{Im}[\mathrm{e}^{ix}]\mathrm{Im}[\mathrm{e}^{iy}]=\cos(x)\cos(y)-sin(x)\sin(y)\\
&\Rightarrow \qquad \cos(x+y)-\cos(x-y)=-2\sin(x)\sin(y)
\end{aligned}
$$

Thus the condition is

$$\underbrace{-\cos(ka)-\cos(Ga)}_{2\sin\left((k+G)a\right)\sin\left((k-G)a\right)}+\frac{mV_0}{\hbar^2 G}\sin(Ga)=0$$

which can be written in the form

$$Z(Ga)=\cos(ka)\qquad\text{with}\qquad Z(x)\stackrel{\text{def}}{=}\cos(x)-\frac{mV_0a}{\hbar^2}\frac{\sin(x)}{x}$$

With relation between G and the energy Eq. 16.9, we obtain

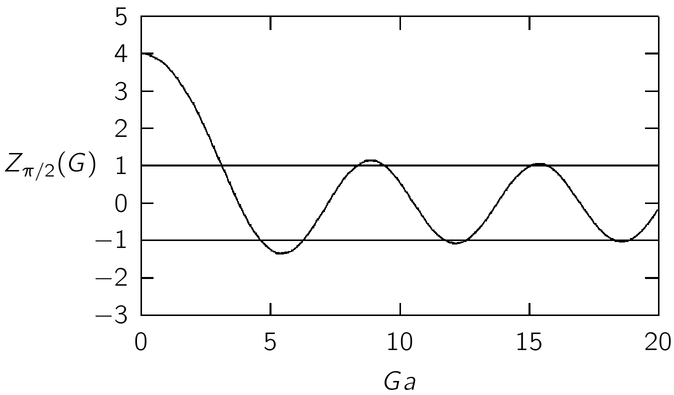
$$k(E)=\frac{1}{a}\arccos\left[Z(\frac{a}{\hbar}\sqrt{2mE})\right] \tag{16.14}$$

It seems to be difficult to resolve the equation for $G(k)$. However it is more straight forward to determine inverse function $k(G)$.

$$k(E)=\frac{1}{a}\arccos\Big(Z(\frac{a}{\hbar}\sqrt{2mE})\Big)$$

Thus we need to determine the zeros of the function $Z_k(G)$.

The figure shows the function $Z_{\pi/2}(Ga)$ for $\frac{mV_0a}{\hbar^2}=5$. We have chosen $k=\pi/2$ because then the $cos(ka)$ drops out. States can only exist where $|Z_k|<1$. Thus we can directly determine the band edges.

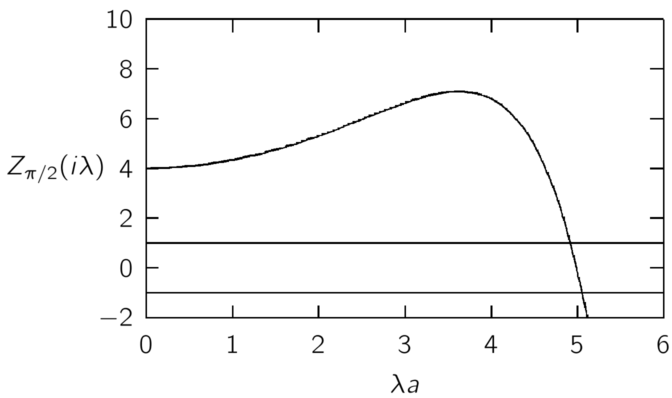


Note that we also need to consider possible states with negative energies! In that case we need to replace $G=i\lambda$ by an imaginary quantity. We use

$$\sin(ix)=\frac{e^{-x}-e^{x}}{2i}=-\frac{1}{i}\sinh(x)=i\sinh(x)$$

The function $Z(i\lambda)$ has the form

$$Z(i\lambda)=\cosh(\lambda a)-\frac{mV_0}{\hbar^2\lambda}\sinh(\lambda a)$$



We see immediately that there is a very narrow band at energy $E = -5$, which corresponds to the bound states for the isolated atom.

```
      PROGRAM MAIN
      IMPLICIT NONE
      INTEGER(4),PARAMETER :: NE=1000       ! #(ENERGY GRID POINTS)
      REAL(8)   ,PARAMETER :: EMIN=-13.D0  ! LOWER BOUND OF ENERGY WINDOW
      REAL(8)   ,PARAMETER :: EMAX=100.D0  ! UPPER BOUND OF ENERGY WINDOW
      REAL(8)   ,PARAMETER :: A=1.D0       ! LATTICE CONSTANT
      REAL(8)   ,PARAMETER :: HBAR=1.D0    ! HBAR
      REAL(8)   ,PARAMETER :: M=1.D0       ! MASS
      REAL(8)   ,PARAMETER :: V=5.D0       !
      INTEGER(4)           :: I
      REAL(8)              :: E,G,Y,Z,K
      REAL(8)              :: PI,G0
!     ****************************************************************
      PI=4.D0*DATAN(1.D0)        ! CALCULATE PI
      G0=2.D0*PI/A               ! RECIPROCAL LATTICE VECTOR
      DO I=1,NE
        E=EMIN+(EMAX-EMIN)/REAL(NE-1)*REAL(I-1)
        IF(E.GT.0) THEN          ! SELECT SCATTERING SOLUTION
          G=SQRT(2.D0*M*E)       ! PSI(X)=A*E^(I*G*X)+B*E^(-I*G*X)+B
          Y=G*A                  ! SCALED VARIABLE
          Z=COS(Y)-M*V*A/(HBAR**2)*SIN(Y)/Y   ! Z(G*A)
        ELSE                     ! SELECT BOUND SOLUTION
          G=SQRT(-2.D0*M*E)      ! PSI(X)=A*E^(I*G*X)+B*E^(-I*G*X)+B
          Y=G*A                  ! SCALED VARIABLE
          Z=COSH(Y)-M*V*A/(HBAR**2)*SINH(Y)/Y  !Z(I*LAMBDA*A)
        END IF
!       == IN THE FORBIDDEN REGION WE SET Z TO THE MAXIMUM OR MINIMUM
!       == ALLOWED VALUE. THE RESULT WILL OCCUR AT THE BOUNDARY OF THE
!       == FIGURE OR ON THE E-AXIS AND WILL BE HIDDEN BEHIND THE FRAME
        Z=MIN(Z,1.D0)            ! SET Z TO MAXIMUM ALLOWED VALUE
        Z=MAX(Z,-1.D0)           ! SET Z TO MINIMUM ALLOWED VALUE
        K=ACOS(Z)/A              ! CALCULATE K(Z)
        IF(K.GT.0.5D0*G0)K=K-G0 ! MAP K INTO FIRST BRILLOUIN ZONE
        PRINT*,E,K/G0,-K/G0      ! PRINT RESULT WITH K IN UNITS OF G0
      ENDDO
      STOP
      END
```

We can now determine the band structure

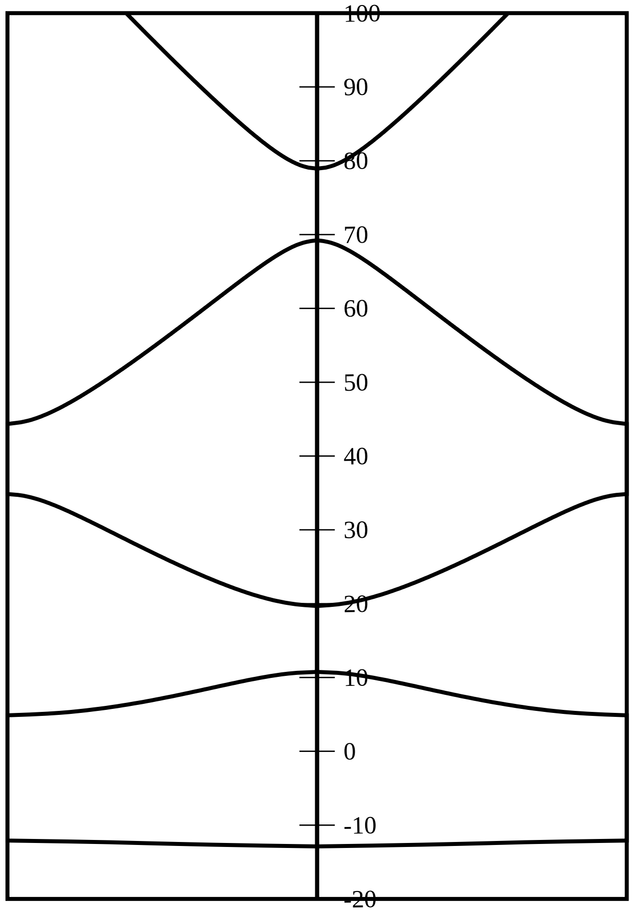


Fig. 16.1: Band structure $E_n(k)$ of the Dirac comb with potential $V_0 = -5$. The horizontal axis ranges from $-g/2$ to $+g/2$, where $g = 2\pi/a$ is the primitive reciprocal lattice vector.The state at negative energies corresponds to the bound states of the atoms, while those at positive energies are scattering states. Compare this dispersion relation with the folded dispersion relation of a free particle (parabola).

16.8 Model for optical absorption

Problem

Consider a on-dimensional particle in a box with light shining on it. The light creates an electric field

$$E(x,t) = E_0 x \left[\theta(-t)\mathrm{e}^{i(\omega - i\lambda)t} + \theta(t)\mathrm{e}^{i(\omega + i\lambda)t}\right]$$

λ describes, how the light is turned on and off again.

Calculate the occupation of the first and second excited state as function of time, if the system starts in the ground state.

Solution:

The box has the length L and extends from $-\frac{L}{2}$ to $\frac{L}{2}$. Inside the box, the first three states have the form

$$\phi_0(x) = \sqrt{\frac{2}{L}}\cos(\frac{\pi x}{L})$$
$$\phi_1(x) = \sqrt{\frac{2}{L}}\sin(\frac{2\pi x}{L})$$
$$\phi_2(x) = \sqrt{\frac{2}{L}}\cos(\frac{3\pi x}{L})$$

Outside of the box the wave functions vanish. Their energies are

$$E_0 = \frac{\hbar^2}{2m}\left(\frac{\pi}{L}\right)^2 = \frac{\hbar^2\pi^2}{2mL^2}$$
$$E_1 = \frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2 = 4\frac{\hbar^2\pi^2}{2mL^2}$$
$$E_2 = \frac{\hbar^2}{2m}\left(\frac{3\pi}{L}\right)^2 = 9\frac{\hbar^2\pi^2}{2mL^2}$$

The time dependent states without perturbation have the form

$$\Psi_j(x,t) = \phi_j(x)\mathrm{e}^{\frac{i}{\hbar}E_j t}$$

Now we use Eq.**??** to evaluate the coefficients.

$$a_m(t) = a_m(-\infty) - \frac{i}{\hbar}\sum_n \int_{-\infty}^{\infty} dt'\ \langle\phi_m|W(t')|\phi_n\rangle \mathrm{e}^{-\frac{i}{\hbar}(E_j-E_n)t'} a_n(t')$$

First we need the coefficients at $t=-\infty$. If the system is in the ground state,

$$a_j(-\infty) = \delta_{j,0}$$

Next we need to evaluate the matrix elements

$$\langle\phi_m|W(t')|\phi_n\rangle = \langle\phi_m|E_0 x|\phi_n\rangle \left[\theta(-t)\mathrm{e}^{i(\omega-i\lambda)t} + \theta(-t)\mathrm{e}^{i(\omega+i\lambda)t}\right]$$

We use

$$\int dx\, \underbrace{x}_{f}\underbrace{\mathrm{e}^{ikx}}_{g'} = \int dx\, \left\{\partial_x\left(\underbrace{x}_{f}\underbrace{\frac{1}{ik}\mathrm{e}^{ikx}}_{g}\right) - \underbrace{1}_{f'}\underbrace{\frac{1}{ik}\mathrm{e}^{ikx}}_{g}\right\} = \frac{x\mathrm{e}^{ikx}}{ik} + \frac{\mathrm{e}^{ikx}}{k^2}$$

to evaluate the formulas needed for the evaluation of the matrix elements

$$
\begin{aligned}
&\int dx\, x\cos(k_1x)\sin(k_2x)\\
&=\int dx\, x\frac{e^{ik_1x}+e^{-ik_1x}}{2}\frac{e^{ik_2x}-e^{-ik_2x}}{2i}\\
&=\frac{1}{4i}\int dx\, x\left(e^{i(k_1+k_2)x}-e^{i(k_1-k_2)x}+e^{-i(k_1-k_2)x}-e^{-i(k_1+k_2)x}\right)\\
&=\frac{-x}{4}\left(\frac{e^{i(k_1+k_2)x}}{k_1+k_2}-\frac{e^{i(k_1-k_2)x}}{k_1-k_2}+\frac{e^{-i(k_1-k_2)x}}{k_1-k_2}-\frac{e^{-i(k_1+k_2)x}}{k_1+k_2}\right)\\
&+\frac{1}{4i}\left(\frac{e^{i(k_1+k_2)x}}{(k_1+k_2)^2}-\frac{e^{i(k_1-k_2)x}}{(k_1-k_2)^2}+\frac{e^{-i(k_1-k_2)x}}{(k_1-k_2)^2}-\frac{e^{-i(k_1+k_2)x}}{(k_1+k_2)^2}\right)\\
&=\frac{-ix}{2}\left(\frac{\sin((k_1+k_2)x)}{k_1+k_2}-\frac{\sin((k_1-k_2)x)}{k_1-k_2}\right)\\
&+\frac{1}{2}\left(\frac{\sin((k_1+k_2)x)}{(k_1+k_2)^2}-\frac{\sin((k_1-k_2)x)}{(k_1-k_2)^2}\right)\\
\Rightarrow\quad &\int_{-a}^{a} dx\, x\cos(k_1x)\sin(k_2x)=\frac{\sin((k_1+k_2)a)}{(k_1+k_2)^2}-\frac{\sin((k_1-k_2)a)}{(k_1-k_2)^2}
\end{aligned}
$$

Now we can evaluate the matrix elements The matrix elements between two identical states vanishes, because the probability density of a state is inversion symmetric, while x is antisymmetric with respect to inversion. Thus the integrand is antisymmetric. Hence the integral vanishes.

$$\langle\phi_i|E_0\hat{x}|\phi_i\rangle=0$$

Similarly the product of ϕ_1 and ϕ_3 is symmetric with respect to inversion, so that the matrix elements between these two states vanishes as well.

$$\langle\phi_1|E_0\hat{x}|\phi_3\rangle=0$$

We still need to evaluate the matrix elements with the second state:

$$
\begin{aligned}
\langle\phi_1|E_0\hat{x}|\phi_2\rangle&=E_0\frac{2}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} dx\;\cos(\frac{\pi}{L}x)\sin(\frac{2\pi}{L}x)\\
&=E_0\frac{2}{L}\left[\frac{\sin(\frac{3\pi}{L}\frac{L}{2})}{\left(\frac{3\pi}{L}\right)^2}-\frac{\sin(-\frac{\pi}{L}\frac{L}{2})}{\left(-\frac{\pi}{L}\right)^2}\right]\\
&=E_0\frac{2}{L}\left(\frac{L}{\pi}\right)^2\left[\frac{\sin(\frac{3\pi}{2})}{9}-\sin(-\frac{\pi}{2})\right]\\
&=\frac{16}{9\pi^2}LE_0
\end{aligned}
$$

and between the second and the third

$$
\begin{aligned}
\langle\phi_3|E_0\hat{x}|\phi_2\rangle&=E_0\frac{2}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} dx\;\cos(\frac{3\pi}{L}x)\sin(\frac{2\pi}{L}x)\\
&=E_0\frac{2}{L}\left[\frac{\sin(\frac{5\pi}{L}\frac{L}{2})}{\left(\frac{5\pi}{L}\right)^2}-\frac{\sin(\frac{\pi}{L}\frac{L}{2})}{\left(\frac{\pi}{L}\right)^2}\right]\\
&=E_0\frac{2}{L}\left(\frac{L}{\pi}\right)^2\left[\frac{\sin(\frac{5\pi}{2})}{25}-\sin(\frac{\pi}{2})\right]\\
&=-\frac{48}{25\pi^2}LE_0
\end{aligned}
$$

Now we insert these matrix elements into the first-order expression

$$
\begin{aligned}
a_2(\infty) &= -\frac{i}{\hbar}\frac{16}{9\pi^2}LE_0\int_{-\infty}^{\infty} dt' \left[\theta(-t')e^{i(\omega-i\lambda)t'} + \theta(t')e^{i(\omega+i\lambda)t'}\right] e^{-\frac{i}{\hbar}(E_2-E_1)t'} \\
&= -\frac{i}{\hbar}\frac{16}{9\pi^2}LE_0\left[\int_{-\infty}^{0} dt'\; e^{\frac{i}{\hbar}(\hbar\omega-(E_2-E_1)-i\hbar\lambda)t} + \int_{0}^{\infty} dt'\; e^{\frac{i}{\hbar}(\hbar\omega-(E_2-E_1)+i\hbar\lambda)t'}\right] \\
&= -\frac{16}{9\pi^2}LE_0\frac{2i\hbar\lambda}{(\hbar\omega-(E_2-E_1))^2+\hbar^2\lambda^2}
\end{aligned}
$$

16.9 Gamow's theory of alpha decay

See Introduction to quantum mechanics by David J. Griffiths

16.10 Transmission coefficient of a tunneling barrier

See Introduction to quantum mechanics by David J. Griffiths

16.11 Fowler-Nordheim tunneling

Editor: not ready!!!

Fowler-Nordheim tunneling is used to determine the band offsets between materials, e.g. between a metal and an oxide.

$$
J_{FN} = C_{FN}E_{ox}^2 \exp\left(-\frac{4}{3}\frac{\sqrt{2m_{ox}^*}}{q\hbar}\frac{(q\phi_B)^{\frac{3}{2}}}{E_{ox}}\right)
$$

Here E_{ox} is the electric field in the oxide, ϕ_B is the barrier height, that is the conduction band offset.

The effective mass of electrons in SiO_2 is $m_{ox}^* = 0.42m_e$

One plots $\ln\left(\frac{J_{FN}}{E_{ox}^2}\right)$ versus $\frac{1}{E_{ox}}$

Fowler nordheim tunneling is applicable when $E_{ox}d > \phi_B$, that is the oxide must be thick enough that the electrons can tunnel into the oxide conduction band.

This formula is derived using the WKB method.

=======================================

Other transport processes are the Pole Frenkel emission. This mechanism dominates if the oxide has a high density of traps. (e.g.Si_3N_4). The current is a simple drift mechanism

$$
\vec{J} = qn\mu\vec{E}
$$

where the carrier density N depends exponentially on the depth of the trap, which is connected to the electric field.

$$
n = n_0 \exp\left[-\frac{q}{k_BT}\left(\phi_B - \sqrt{\frac{qE}{\pi\epsilon_N}}\right)\right]
$$

so that the total current equals

$$
\vec{J} = qn_0 \exp\left[-\frac{q}{k_BT}\left(\phi_B - \sqrt{\frac{qE}{\pi\epsilon_N}}\right)\right]\mu\vec{E}
$$

============================== Space charge limited current see www B. Van Zeghbroeck.

Part I

Appendix

Appendix A

Galilei invariance

Here we show, that the dynamics of the envelope function $\chi(x,t)$ of a wave packet is Galilei invariant, respectively that it is invariant under the transformation from a resting to a moving reference frame.

For the derivation it is important to distinguish between partial and total derivatives. In order to make this explicit we consider

$$\chi_v(x,t) \stackrel{\text{def}}{=} \chi_0(x-vt,t) = \chi_0(y(x,t),\tau(x,t)) \tag{A.1}$$

with the transformed variables

$$y(x,t) \stackrel{\text{def}}{=} x' = x - vt \qquad \text{and} \qquad \tau(x,t) \stackrel{\text{def}}{=} t' = t$$

where $\chi_0(x,t)$ is the resting wave packet, and $\chi_v(x,t)$ is the moving wave packet.

Now, we derive an equation for the envelope function $\chi_v(x,t)$ from the Schrödinger equation for the moving wave packet

$$\phi_v(x,t) = \chi_v(x,t)\mathrm{e}^{i(kx-\omega t)} \tag{A.2}$$

The following expressions will be used:

$$\begin{aligned}
i\hbar\partial_t \mathrm{e}^{i(kx-\omega t)}\chi_v(x,t) &= \mathrm{e}^{i(kx-\omega t)}[\hbar\omega + i\hbar\partial_t]\chi_v(x,t)\\
\frac{\hbar}{i}\partial_x \mathrm{e}^{i(kx-\omega t)}\chi_v(x,t) &= \mathrm{e}^{i(kx-\omega t)}[\hbar k + \frac{\hbar}{i}\partial_x]\chi_v(x,t)\\
\left(\frac{\hbar}{i}\partial_x\right)^2 \mathrm{e}^{i(kx-\omega t)}\chi_v(x,t) &= \mathrm{e}^{i(kx-\omega t)}\left[\hbar k + \frac{\hbar}{i}\partial_x\right]^2\chi_v(x,t)
\end{aligned}$$

From the Schrödinger equation for the moving wave packet we derive a differential equation for its envelope function $\chi_v(x,t)$ alone.

$$0 = \left(i\hbar\partial_t + \frac{\hbar^2}{2m}\partial_x\right)\phi_v(x,t) \tag{A.3}$$

$$\stackrel{Eq.\ A.2}{\Rightarrow} \qquad 0 = \left(i\hbar\partial_t + \frac{\hbar^2}{2m}\partial_x\right)\mathrm{e}^{i(kx-\omega t)}\chi_v(x,t) = \mathrm{e}^{i(kx-\omega t)}\left[\hbar\omega + i\hbar\partial_t - \frac{1}{2m}\left(\hbar k + \frac{\hbar}{i}\partial_x\right)^2\right]\chi_v(x,t) \tag{A.4}$$

$$\Rightarrow \qquad 0 = \left[\hbar\omega + i\hbar\partial_t - \frac{1}{2m}\left(\hbar k + \frac{\hbar}{i}\partial_x\right)^2\right]\chi_v(x,t) \tag{A.5}$$

Now we can rewrite

$$
\begin{aligned}
i\hbar\partial_t\chi_v(x,t) &\stackrel{\text{Eq. }A.1}{=} i\hbar\partial_t\chi_0(y(x,t),\tau(x,t)) \\
&= i\hbar\partial_y\chi_0(y(x,t),\tau(x,t))\underbrace{\frac{\partial y}{\partial_t}}_{-v}+i\hbar\partial_\tau\chi_0(y(x,t),\tau(x,t))\underbrace{\frac{\partial\tau}{\partial_t}}_{1} \\
&= -vi\hbar\partial_y\chi_0(y,\tau)+i\hbar\partial_\tau\chi_0(y,\tau) \\
\frac{\hbar}{i}\partial_x\chi_v(x,t) &\stackrel{\text{Eq. }A.1}{=} \frac{\hbar}{i}\partial_x\chi_0(y(x,t),\tau(x,t)) \\
&= \frac{\hbar}{i}\partial_y\chi_0(y(x,t),\tau(x,t))\underbrace{\frac{\partial y}{\partial_x}}_{1}+\frac{\hbar}{i}\partial_\tau\chi_0(y(x,t),\tau(x,t))\underbrace{\frac{\partial\tau}{\partial_x}}_{0} \\
&= \frac{\hbar}{i}\partial_y\chi_0(y,\tau)
\end{aligned}
$$

With this we can turn the equation for the envelope function χ_v into an equation for the envelope function $\chi_0(x,t)$ for the resting system.

$$
\begin{aligned}
0 &\stackrel{\text{Eq. }A.5}{=} \left[\hbar\omega+i\hbar\partial_t-\frac{1}{2m}\left(\hbar k+\frac{\hbar}{i}\partial_x\right)^2\right]\underbrace{\chi_0(y(x,t),\tau(x,t))}_{\chi_v(x,t)} \\
&= \left[\hbar\omega-vi\hbar\partial_y+i\hbar\partial_\tau-\frac{1}{2m}\left(\hbar k+\frac{\hbar}{i}\partial_y\right)^2\right]\chi_0(y,\tau) \\
&= \left[\hbar\omega-vi\hbar\partial_y+i\hbar\partial_\tau-\frac{(\hbar k)^2}{2m}-\frac{\hbar k}{m}\frac{\hbar}{i}\partial_y-\frac{-\hbar^2}{2m}\partial_y^2\right]\chi_0(y,\tau) \\
&= \left[i\hbar\partial_\tau+\frac{\hbar^2}{2m}\partial_y^2+\left(\hbar\omega-\frac{(\hbar k)^2}{2m}\right)+\left(v-\frac{\hbar k}{m}\right)\frac{\hbar}{i}\partial_y\right]\chi_0(y,\tau)
\end{aligned}
$$

If $\hbar\omega=E(k)=\frac{\hbar^2k^2}{2m}$ and $v=\frac{\partial E}{\partial(\hbar k)}$, we obtain the Schrödinger equation itself

$$
\left[i\hbar\partial_\tau+\frac{\hbar^2}{2m}\partial_y^2\right]\chi_0(y,\tau)
$$

Thus the transformation into a moving reference frame involves a simple transformation of the envelope function

$$
\underbrace{x'}_{y}=x-vt \qquad\text{and}\qquad \underbrace{t'}_{\tau}=t \qquad\text{and}\qquad \underbrace{\chi'(x',t')}_{\chi_0(y,\tau)}=\underbrace{\chi(x,t)}_{\chi_0(x,t)}
$$

Thus a wave packet at rest $\chi_0(x,t)$ is transformed into a moving one by firstly replacing the spatial argument by $x-vt$ and secondly by multiplication with a plane wave $\mathrm{e}^{i(kx-\omega(k)t)}$, where $v=\partial\omega/\partial k$.

$$
\begin{aligned}
\psi'(x,t) &\stackrel{\text{Eq. }A.2}{=} \chi_v(x,t)\mathrm{e}^{i(kx-\omega(k)t)}=\chi_0(x-vt,t)\mathrm{e}^{i(kx-\omega(k)t)} \\
&\stackrel{\text{Eq. }A.2}{=} =\psi(x-vt,t)\mathrm{e}^{i(kx-\omega(k)t)}
\end{aligned}
$$

In order to transform the wave function itself, we need to add a phase factor. The reason is the following: The envelope function only describes the spatial probability distribution. The information about the velocities is not contained in the envelope function. This information must be included in the transformation of the wave function. The phase factor tells that the velocity distribution of the wave packet has to be shifted by the relative velocity v.

Appendix B

Spreading wave packet of a free particle

Let us explore a one-dimensional free particle with a wave packet that initially has Gaussian shape. The Schrödinger equation has the form

$$i\hbar\partial_t\psi(x,t) = \frac{-\hbar^2}{2m}\partial_x^2\psi(x,t) \tag{B.1}$$

We use an Ansatz for the wave function that has the same form as the initial wave packet, but with time dependent coefficients

$$\psi(x,t) = A(t)\mathrm{e}^{-\frac{1}{2}B(t)x^2} \tag{B.2}$$

The parameters A, B and their time dependence are obtained by insertion into the time-dependent Schrödinger equation on the one side and the normalization condition on the other side. [1]

Let us first evaluate the derivatives that appear in the Schrödinger equation Eq. B.1.

$$i\hbar\partial_t\psi(x,t) \stackrel{\text{Eq. }B.2}{=} i\hbar\partial_t\Big[A(t)\mathrm{e}^{-\frac{1}{2}B(t)x^2}\Big] = i\hbar\Big[\frac{\partial_t A}{A} - \frac{1}{2}(\partial_t B)x^2\Big]\Big[A(t)\mathrm{e}^{-\frac{1}{2}B(t)x^2}\Big] \tag{B.3}$$

$$\begin{aligned}\frac{-\hbar^2}{2m}\partial_x^2\psi(x,t) &\stackrel{\text{Eq. }B.2}{=} \frac{-\hbar^2}{2m}\partial_x^2\Big[A(t)\mathrm{e}^{-\frac{1}{2}B(t)x^2}\Big] = \frac{-\hbar^2}{2m}A(t)\partial_x\Big[-B(t)x\mathrm{e}^{-\frac{1}{2}B(t)x^2}\Big] \\ &= \frac{-\hbar^2}{2m}\Big[-B(t)+B^2(t)x^2\Big]A(t)\mathrm{e}^{-\frac{1}{2}B(t)x^2}\end{aligned} \tag{B.4}$$

From the requirement that the left-hand side, Eq. B.3, and the right-hand side, Eq. B.4, of the Schrödinger equation Eq. B.1, are identical everywhere in space, we obtain the two conditions for $A(t)$ and $B(t)$

$$i\hbar\Big(\frac{\partial_t A}{A}\Big) = \frac{-\hbar^2}{2m}\Big(-B(t)\Big) \tag{B.5}$$

$$i\hbar\Big(-\frac{1}{2}\partial_t B\Big) = \frac{-\hbar^2}{2m}\Big(B^2(t)\Big) \tag{B.6}$$

This is a system of ordinary differential equations that can be solved as follows: We begin with

[1] The factor $\frac{1}{2}$ has been included only because the formulas for Gaussians with this choice turn out to be more pleasant.

the second equation for $B(t)$

$$\begin{aligned}
\partial_t B(t) &\stackrel{\text{Eq. } B.6}{=} \frac{-i\hbar}{m} B^2(t) &\text{(B.7)}\\
\int_0^t dt\ B^{-2}(t)\partial_t B(t) &= \int_0^t dt\ \frac{-i\hbar}{m} &\text{(B.8)}\\
-\frac{1}{B(t)} + \frac{1}{B(0)} &= \frac{-i\hbar}{m} t &\text{(B.9)}\\
\Rightarrow \qquad \partial_t \frac{1}{B(t)} = -\frac{\partial_t B(t)}{B^2(t)} &\stackrel{\text{Eq. } B.7}{=} \frac{i\hbar}{m}\\
\Rightarrow \qquad \frac{1}{B(t)} &= C_0 + \int_0^t dt'\ \frac{i\hbar}{m} = \underbrace{C_0 + \frac{i\hbar}{m} t}_{\stackrel{\text{def}}{=} C(t)}\\
\Rightarrow \qquad B(t) &\stackrel{\text{Eq. } B.11}{=} \frac{1}{C(t)} &\text{(B.10)}
\end{aligned}$$

where we introduced the integration constant C_0 and the variable

$$C(t) \stackrel{\text{def}}{=} C_0 + \frac{i\hbar t}{m} \tag{B.11}$$

to simplify the following equations

Next we insert this result, Eq. B.10, into the first equation, Eq. B.5, to obtain $A(t)$

$$\begin{aligned}
A^{-1}\partial_t A(t) &= \partial_t \ln[A(t)] \stackrel{\text{Eq. } B.5}{=} \frac{-i\hbar}{2m} B(t)\\
\Rightarrow \qquad \ln[A(t)] &= \ln[A(0)] + \frac{-i\hbar}{2m}\int_0^t dt'\ B(t') \stackrel{\text{Eq. } B.10}{=} \ln[A(0)] + \frac{-i\hbar}{2m}\int^t dt'\ \frac{1}{C(t')}\\
&\stackrel{\text{Eq. } B.11}{=} \ln[A(0)] - \frac{1}{2}\int_0^t dt'\ \frac{dC(t')}{dt'}\frac{1}{C(t')} = \ln[A(0)] - \frac{1}{2}\int_0^t dt'\ \frac{d}{dt'}\ln[C(t')]\\
&= \ln[A(0)] - \frac{1}{2}\ln[C(t)] + \frac{1}{2}\ln[C_0] \stackrel{\text{Eq. } B.13}{=} D - \frac{1}{2}\ln[C(t)]\\
\Rightarrow \qquad A(t) &= \frac{1}{\sqrt{C(t)}}\mathrm{e}^D &\text{(B.12)}
\end{aligned}$$

where

$$D \stackrel{\text{def}}{=} \ln[A(0)] + \frac{1}{2}\ln[C_0] \tag{B.13}$$

Thus the wave function has the form

$$\psi(x,t) \stackrel{Eqs.\ B.2,B.10,B.12}{=} \mathrm{e}^D \frac{1}{\sqrt{C(t)}}\mathrm{e}^{-\frac{1}{2}\frac{x^2}{C(t)}} \tag{B.14}$$

Finally we fix the factor e^D using[2] the normalization condition $\int dx\ \psi^*\psi$.

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} dx\ \psi^*(x,t)\psi(x,t) = \frac{\mathrm{e}^{D+D^*}}{|C(t)|}\int_{-\infty}^{\infty} dx\ \mathrm{e}^{-\frac{1}{2}\left(\frac{1}{C(t)^*}+\frac{1}{C(t)}\right)x^2}\\
&= \frac{\mathrm{e}^{2\mathrm{Re}[D]}}{|C(t)|}\int_{-\infty}^{\infty} dx\ \mathrm{e}^{-\frac{1}{2}\frac{2\mathrm{Re}[C(t)]}{|C(t)|^2}x^2} = \mathrm{e}^{2\mathrm{Re}[D]}\frac{1}{|C(t)|}\sqrt{\frac{|C(t)|^2}{2\mathrm{Re}[C(t)]}}\sqrt{2\pi}\\
&= \mathrm{e}^{2\mathrm{Re}[D]}\sqrt{\frac{\pi}{\mathrm{Re}[C(t)]}} \stackrel{\text{Eq. } B.11}{=} \mathrm{e}^{2\mathrm{Re}[D]}\sqrt{\frac{\pi}{C_0}}\\
\Rightarrow \qquad \mathrm{e}^{\mathrm{Re}[D]} &= \sqrt[4]{\frac{C_0}{\pi}} &\text{(B.15)}
\end{aligned}$$

[2]We also use the **Gauss integral** $\int_{-\infty}^{\infty} dx\ \mathrm{e}^{-\frac{1}{2}x^2} = \sqrt{2\pi}$

The imaginary part of D is irrelevant because that only changes the overall phase of the wave function.

We obtain the wave function

$$\psi(x,t) \stackrel{\text{Eqs. } B.14, B.15}{=} \sqrt[4]{\frac{C_0}{\pi|C(t)|^2}}\,\mathrm{e}^{-\frac{1}{2}\frac{x^2}{C(t)}} \tag{B.16}$$

with $C(t)$ given by Eq. B.11.

B.1 Probability distribution

Let us now evaluate the probability distribution $P(r) = \psi^*(x,t)\psi(x,t)$.

$$\begin{aligned}
P(r,t) = \psi^*(x,t)\psi(x,t) &\stackrel{\text{Eq. } B.16}{=} \sqrt{\frac{C_0}{\pi|C(t)|^2}}\mathrm{e}^{-\frac{1}{2}\left(\frac{1}{C(t)}+\frac{1}{C^*(t)}\right)x^2} \\
&= \sqrt{\frac{C_0}{\pi|C(t)|^2}}\,\mathrm{e}^{-\frac{\mathrm{Re}[C(t)]}{|C(t)|^2}x^2} = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{2C_0}{|C(t)|^2}}\,\mathrm{e}^{-\frac{1}{2}\frac{2C_0}{|C(t)|^2}x^2} \\
&= \frac{1}{\sqrt{2\pi}\cdot d(t)}\mathrm{e}^{-\frac{1}{2}\left(\frac{x}{d(t)}\right)^2}
\end{aligned} \tag{B.17}$$

with

$$d(t) \stackrel{\text{def}}{=} \sqrt{\frac{|C(t)|^2}{2C_0}} \stackrel{\text{Eq. } B.11}{=} \sqrt{\frac{C_0}{2}}\sqrt{1+\left(\frac{\hbar t}{mC_0}\right)^2} = d(0)\sqrt{1+\left(\frac{\hbar t}{2md(0)^2}\right)^2} \tag{B.18}$$

We see that the density maintains its Gaussian shape, but the width $d(t)$ of the Gaussian increases after a certain while almost linearly with time.

The spreading of a Gaussian probability distribution reminds of the behavior of a diffusing particle. Such a particle undergoes **Brownian motion**. The related probability distribution behaves similarly to that of a quantum particle, but the width of the distribution grows only like the square root of time.

We define the width of the wave packet as $\Delta \stackrel{\text{def}}{=} \sqrt{\langle x^2\rangle - \langle x\rangle^2}$. Now we investigate how the width changes with time.

Let us consider the first few **moments**[3] of the probability distribution.

$$\begin{aligned}
\langle 1\rangle &= \int_{-\infty}^{\infty} dx\; P(x,t) \stackrel{\text{Eq. } B.17}{=} 1 \\
\langle x\rangle &= \int_{-\infty}^{\infty} dx\; P(x,t)x \stackrel{\text{Eq. } B.17}{=} 0 \\
\langle x^2\rangle &= \int_{-\infty}^{\infty} dx\; P(x,t)x^2 \stackrel{\text{Eq. } B.17}{=} \frac{1}{\sqrt{2\pi}\cdot d(t)}\int_{-\infty}^{\infty} dx\;\mathrm{e}^{-\frac{1}{2}\left(\frac{x}{d(t)}\right)^2}x^2 \\
&\stackrel{u=x/(\sqrt{2}d)}{=} \frac{2d(t)^2}{\sqrt{\pi}}\int_{-\infty}^{\infty} du\;\mathrm{e}^{-u^2}u^2 = \frac{2d(t)^2}{\sqrt{\pi}}\frac{\sqrt{\pi}}{2} = d(t)^2
\end{aligned} \tag{B.19}$$

This reveals the physical meaning of the variable $d(t)$.

[3] The nth moment of a real function $f(x)$ is defined as

$$\langle x^n\rangle \stackrel{\text{def}}{=} \int dx\; f(x)x^n$$

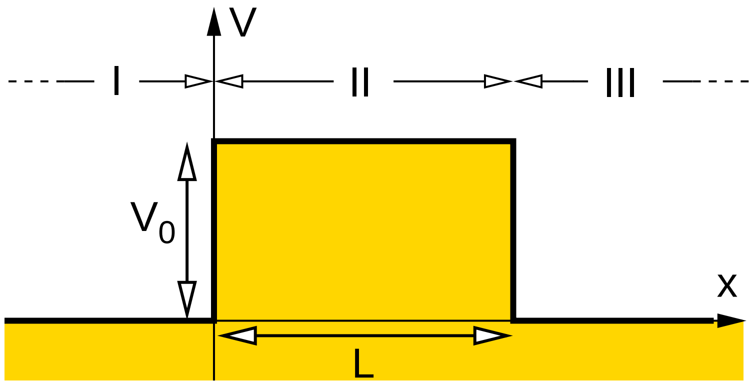
We combine Eq. B.19 to obtain the width

$$\Delta(t) \stackrel{\text{def}}{=} \sqrt{\left\langle (x-\langle x\rangle)^2 \right\rangle} = \sqrt{\langle x^2\rangle - \langle x\rangle^2} \stackrel{\text{Eq. }B.19}{=} d(t)$$
$$\stackrel{\text{Eq. }B.18}{=} d(0)\sqrt{1+\left(\frac{\hbar t}{2md^2(0)}\right)^2} \rightarrow \frac{\hbar t}{2md^2(0)} \quad \text{for t} \rightarrow \infty$$

Thus the width of the probability distribution grows linearly with time.

Appendix C

The one-dimensional rectangular barrier



The potential of a one-dimensional rectangular barrier has the following form

$$V(x) = \begin{Bmatrix} V_0 \\ 0 \end{Bmatrix} \text{for} \begin{Bmatrix} 0 < x < L \\ x < 0; x > L \end{Bmatrix}$$

We call the region to the left of the barrier region I, the barrier itself region II and what is to the right of the barrier is region III.

The wave number of the particle in region I and III is

$$k_I = \frac{1}{\hbar}\sqrt{2mE} \tag{C.1}$$

and the wave number in region II is

$$k_{II} = \frac{1}{\hbar}\sqrt{2m(E-V)} \tag{C.2}$$

At first we consider the case where the energy of the particle is larger than the barrier. Later we will use this result to obtain the results for particles with an energy that is lower than the barrier.

C.1 $E > V$, the barrier can be classically surmounted

We start with the Ansatz

$$
\begin{aligned}
\psi_I(x) &= A_I \mathrm{e}^{ik_I x} + B_I \mathrm{e}^{-ik_I x} \\
\psi_{II}(x) &= A_{II} \mathrm{e}^{ik_{II} x} + B_{II} \mathrm{e}^{-ik_{II} x} \\
\psi_{III}(x) &= A_{III} \mathrm{e}^{ik_I x} + B_{III} \mathrm{e}^{-ik_I x}
\end{aligned}
$$

We select one particular partial solution, namely the one corresponding to a particle beam approaching the barrier from the left, meaning that no particles arrive from the right-hand side or

$$B_{III} = 0 \tag{C.3}$$

A second coefficient can be fixed by the choosing a common prefactor. We choose the prefactor so that

$$A_I = 1 \tag{C.4}$$

The remaining four coefficients are determined by the requirement that the wave function is continuous and differentiable, i.e.

$$\psi_I(0) = \psi_{II}(0) \tag{C.5}$$
$$\partial_x \psi_I(0) = \partial_x \psi_{II}(0) \tag{C.6}$$
$$\psi_{II}(L) = \psi_{III}(L) \tag{C.7}$$
$$\partial_x \psi_{II}(L) = \partial_x \psi_{III}(L) \tag{C.8}$$

These requirements translate into

$$
\begin{aligned}
1 + B_I &\stackrel{Eqs.\ C.5,C.4}{=} A_{II} + B_{II} \\
ik_I - ik_I B_I &\stackrel{Eqs.\ C.6,C.4}{=} ik_{II} A_{II} - ik_{II} B_{II} \\
A_{II} \mathrm{e}^{ik_{II} L} + B_{II} \mathrm{e}^{-ik_{II} L} &\stackrel{Eqs.\ C.7,C.3}{=} A_{III} \mathrm{e}^{ik_I L} \\
ik_{II} A_{II} \mathrm{e}^{ik_{II} L} - ik_{II} B_{II} \mathrm{e}^{-ik_{II} L} &\stackrel{Eqs.\ C.8,C.3}{=} ik_I A_{III} \mathrm{e}^{ik_I L}
\end{aligned}
$$

We simplify the equations

$$1 + B_I = A_{II} + B_{II} \tag{C.9}$$
$$\frac{k_I}{k_{II}}(1 - B_I) = A_{II} - B_{II} \tag{C.10}$$
$$A_{II} \mathrm{e}^{-i(k_I - k_{II})L} + B_{II} \mathrm{e}^{-i(k_I + k_{II})L} = A_{III} \tag{C.11}$$
$$A_{II} \mathrm{e}^{-i(k_I - k_{II})L} - B_{II} \mathrm{e}^{-i(k_I + k_{II})L} = \frac{k_I}{k_{II}} A_{III} \tag{C.12}$$

By adding and subtracting the last two equations Eq. C.11 and Eq. C.12 we can express A_{II} and B_{II} by A_{III}:

$$A_{II} \stackrel{Eqs.\ C.11,C.12}{=} \frac{1}{2} \mathrm{e}^{i(k_I - k_{II})L} \left(1 + \frac{k_I}{k_{II}}\right) A_{III} \tag{C.13}$$
$$B_{II} \stackrel{Eqs.\ C.11,C.12}{=} \frac{1}{2} \mathrm{e}^{i(k_I + k_{II})L} \left(1 - \frac{k_I}{k_{II}}\right) A_{III} \tag{C.14}$$

Inserting this result, Eq. C.13 and Eq. C.14, into the first equation Eq. C.9 we obtain

$$\begin{aligned}
1+B_I &\overset{\text{Eq. }C.9}{=} (A_{II}+B_{II})\\
&\overset{Eqs.\ C.13,C.14}{=} \frac{1}{2}\mathrm{e}^{i(k_I-k_{II})L}\left(1+\frac{k_I}{k_{II}}\right)A_{III}+\frac{1}{2}\mathrm{e}^{i(k_I+k_{II})L}\left(1-\frac{k_I}{k_{II}}\right)A_{III}\\
&= \left[\frac{\mathrm{e}^{i(k_I-k_{II})L}+\mathrm{e}^{i(k_I+k_{II})L}}{2}+\frac{k_I}{k_{II}}\frac{\mathrm{e}^{i(k_I-k_{II})L}-\mathrm{e}^{i(k_I+k_{II})L}}{2}\right]A_{III}\\
&= \mathrm{e}^{ik_IL}\left[\cos(k_{II}L)-i\frac{k_I}{k_{II}}\sin(k_{II}L)\right]A_{III}
\end{aligned}\tag{C.15}$$

The second equation, Eq. C.10, together with the expressions for A_{II} and B_{II}, Eq. C.13 and Eq. C.14, yields

$$\begin{aligned}
1-B_I &= \frac{k_{II}}{k_I}(A_{II}-B_{II})\\
&= \frac{k_{II}}{k_I}\left[\frac{1}{2}\mathrm{e}^{i(k_I-k_{II})L}\left(1+\frac{k_I}{k_{II}}\right)A_{III}-\frac{1}{2}\mathrm{e}^{i(k_I+k_{II})L}\left(1-\frac{k_I}{k_{II}}\right)A_{III}\right]\\
&= \mathrm{e}^{ik_IL}\left[\frac{1}{2}\mathrm{e}^{-ik_{II}L}\left(1+\frac{k_{II}}{k_I}\right)+\frac{1}{2}\mathrm{e}^{ik_{II}L}\left(1-\frac{k_{II}}{k_I}\right)\right]A_{III}\\
&= \mathrm{e}^{ik_IL}\left[\cos(k_{II}L)-i\frac{k_{II}}{k_I}\sin(k_{II}L)\right]A_{III}
\end{aligned}\tag{C.16}$$

Finally we add Eqs. C.15,C.16 to obtain A_{III}:

$$\begin{aligned}
2 &= \mathrm{e}^{ik_IL}\left[\cos(k_{II}L)-i\frac{k_I}{k_{II}}\sin(k_{II}L)\right]A_{III}+\mathrm{e}^{ik_IL}\left[\cos(k_{II}L)-i\frac{k_{II}}{k_I}\sin(k_{II}L)\right]A_{III}\\
&= \mathrm{e}^{ik_IL}\left[\cos(k_{II}L)-i\frac{k_I}{k_{II}}\sin(k_{II}L)+\cos(k_{II}L)-i\frac{k_{II}}{k_I}\sin(k_{II}L)\right]A_{III}\\
&= \mathrm{e}^{ik_IL}\left[2\cos(k_{II}L)-i\left(\frac{k_I}{k_{II}}+\frac{k_{II}}{k_I}\right)\sin(k_{II}L)\right]A_{III}\\
\Rightarrow A_{III} &= \mathrm{e}^{-ik_IL}\left[\cos(k_{II}L)-\frac{i}{2}\frac{k_I^2+k_{II}^2}{k_Ik_{II}}\sin(k_{II}L)\right]^{-1}
\end{aligned}\tag{C.17}$$

Knowing A_{III}, we can determine the other parameters from Eqs. C.15,C.13, and Eq. C.14. The transmission coefficient T of the barrier can be evaluated as

$$\begin{aligned}
T &= \left|\frac{A_{III}}{A_I}\right|^2\\
&\overset{\text{Eq. }C.17}{=} \left[\left(\cos(k_{II}L)-\frac{i}{2}\frac{k_I^2+k_{II}^2}{k_Ik_{II}}\sin(k_{II}L)\right)\left(\cos(k_{II}L)+\frac{i}{2}\frac{k_I^2+k_{II}^2}{k_Ik_{II}}\sin(k_{II}L)\right)\right]^{-1}\\
&= \left[\underbrace{\cos^2(k_{II}L)}_{1-\sin^2(k_{II}L)}+\left(\frac{k_I^2+k_{II}^2}{2k_Ik_{II}}\right)^2\sin^2(k_{II}L)\right]^{-1}\\
&= \left[1+\left(\left(\frac{k_I^2+k_{II}^2}{2k_Ik_{II}}\right)^2-1\right)\sin^2(k_{II}L)\right]^{-1}
\end{aligned}\tag{C.18}$$

Now we express the wave vectors k_I and k_{II} by the energy [1]

$$
\begin{aligned}
T &\stackrel{\text{Eq. } C.18}{=} \left[1+\left(\left(\frac{k_I^2+k_{II}^2}{2k_Ik_{II}}\right)^2-1\right)\sin^2(k_{II}L)\right]^{-1} \\
&= \left[1+\frac{V^2}{4E(E-V)}\sin^2(\frac{L}{\hbar}\sqrt{2m(E-V)})\right]^{-1}
\end{aligned}
\tag{C.19}
$$

In this derivation we assumed that the energy of the particle is greater than the barrier, so that the motion over the barrier is classically allowed and the wave function is oscillatory also within the barrier.

C.2 Tunneling effect $E < V$

Now we consider the case when the energy is lower than the barrier.

In that case the wave vector within the barrier becomes imaginary, which says, that the wave function is exponentially decaying or exponentially increasing. There is nothing special about a imaginary wave vector, but for convenience we introduce a new quantity, which is real: The decay factor λ.

$$
k_{II} = \frac{1}{\hbar}\sqrt{2m(E-V)} =: i\lambda
$$

We just proceed with the derivation given for $E > V$ and make sure that no operation depends on the fact that the wave vector k_{II} is real.

The only complication is when we introduce the cosinus and the sinus.

$$
\begin{aligned}
\cos(k_{II}L) &= \frac{e^{ik_{II}L}+e^{ik_{II}L}}{2} = \frac{e^{-\lambda L}+e^{+\lambda L}}{2} = \cosh(\lambda L) \\
\sin(k_{II}L) &= \frac{e^{ik_{II}L}-e^{ik_{II}L}}{2i} = \frac{e^{-\lambda L}-e^{+\lambda L}}{2i} = i\sinh(\lambda L)
\end{aligned}
$$

Thus we have to do this replacement in the final expression. We obtain for the transmission coefficient[2]

$$
\begin{aligned}
T &= \left[1-\left(\left(\frac{k_I^2-\lambda^2}{2ik_I\lambda}\right)^2-1\right)\sinh^2(\lambda L)\right]^{-1} \\
&= \left[1+\frac{V^2}{4E(V-E)}\sinh^2(\frac{1}{\hbar}\sqrt{2m(V-E)}L)\right]^{-1}
\end{aligned}
\tag{C.20}
$$

[1] We use the following, where $k_I = \frac{1}{\hbar}\sqrt{2mE}$ and $k_{II} = \frac{1}{\hbar}\sqrt{2m(E-V)}$:

$$
\begin{aligned}
\left(\frac{k_I^2+k_{II}^2}{2k_Ik_{II}}\right)^2-1 &= \left(\frac{2E-V}{2\sqrt{E(E-V)}}\right)^2-1 = \frac{4E^2-4EV+V^2}{4E(E-V)}-1 \\
&= \frac{4E^2-4EV+V^2-4E^2+4EV}{4E(E-V)} = \frac{V^2}{4E(E-V)}
\end{aligned}
$$

[2]

$$
\begin{aligned}
\left(\frac{k_I^2-\lambda^2}{2ik_I\lambda}\right)^2-1 &= \left(\frac{E-(V-E)}{2i\sqrt{E(V-E)}}\right)^2-1 = \frac{4E^2-4EV+V^2}{-4E(V-E)}-1 \\
&= \frac{4E^2-4EV+V^2+4EV-4E^2}{-4E(V-E)} = -\frac{V^2}{4E(V-E)}
\end{aligned}
$$

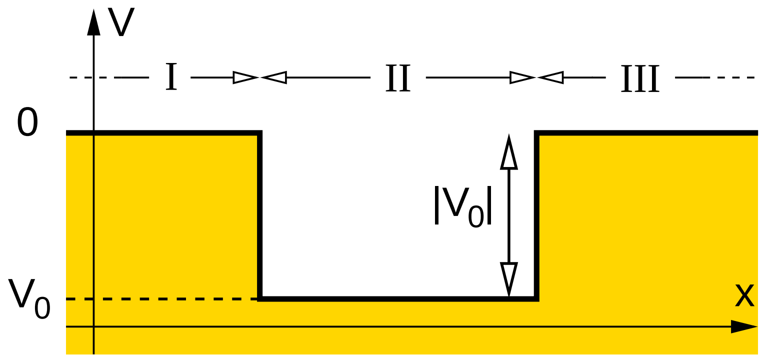
If the barrier is very wide or the barrier is much higher than the energy of the particle, we can simplify the expression to[3]

$$T \approx \frac{4E(V-E)}{V^2} \mathrm{e}^{-\frac{1}{\hbar}\sqrt{2m(V-E)}L}$$

[3]In this limit, the term with the sinus hyperbolicus will always be much larger than one. so that the latter can be neglected. Secondly for large arguments the sinus hyperbolicus will be dominated by one of the two exponential functions, so that we can ignore the other. These arguments sound hand waiving, but root firmly on the ground of a Taylor expansion.

Appendix D

Particle scattering at a one-dimensional square well



D.1 Square well $E > 0$

The barrier can be converted into a well discussed in Appendix by changing the sign of the barrier height.

Thus, for particles that can travel classically outside the well, we can directly use the transmission coefficients of the barrier Eq. C.19.

D.2 Square well $E < 0$

For the energies for which a classical particle is confined to the well, we use the technique of the logarithmic derivatives from Section 4.6.4 on p. 75.

The wave function consists of the three parts: Inside the well we have travelling waves, but outside we have functions that exponentially decay away from the well.

$$\psi_I(x) = A_I e^{\lambda x}$$
$$\psi_{II}(x) = A_{II} e^{ik_{II}x} + B_{II} e^{-ik_{II}x}$$
$$\psi_{III}(x) = B_{III} e^{-\lambda(x-L)}$$

We have to solve two subsequent matching problems. We start from the left and match the solutions ψ_{II} to ψ_I. In this case, we assign the following values to the expressions $\phi_i(x)$ in Eq. 4.44.

$$\begin{aligned}
\psi_1(x) &= \mathrm{e}^{ik_{II}x} & D_1 &= ik_{II} \\
\psi_2(x) &= \mathrm{e}^{-ik_{II}x} & D_2 &= -ik_{II} \\
\psi_3(x) &= \mathrm{e}^{\lambda x} & D_3 &= \lambda
\end{aligned}$$

Insertion into Eq. 4.44 yields

$$\begin{aligned}
\psi_{II}(x) &= A_I \left[\mathrm{e}^{ik_{II}x}\frac{-ik_{II}-\lambda}{-ik_{II}-ik_{II}} + \mathrm{e}^{-ik_{II}x}\frac{ik_{II}-\lambda}{ik_{II}+ik_{II}}\right] \\
&= A_I \left[\mathrm{e}^{ik_{II}x}\left(\frac{1}{2}-i\frac{\lambda}{2k_{II}}\right) + \mathrm{e}^{-ik_{II}x}\left(\frac{1}{2}+i\frac{\lambda}{2k_{II}}\right)\right] \\
&= A_I \left[\cos(k_{II}x) + \frac{\lambda}{k_{II}}\sin(k_{II}x)\right]
\end{aligned}$$

Now we come to the second matching problem at $x = L$. The solution, we just obtained for ψ_{II}, plays now the role of ψ_3, to which we match an exponentially decaying and an exponentially increasing function in region III.

$$\begin{aligned}
\psi_1(x) &= \mathrm{e}^{\lambda x} & D_1 &= \lambda \\
\psi_2(x) &= \mathrm{e}^{-\lambda x} & D_2 &= -\lambda \\
\psi_3(x) = A_I \left[\cos(k_{II}x) + \frac{\lambda}{k_{II}}\sin(k_{II}x)\right] & & D_3 &= \frac{-k_{II}\sin(k_{II}L) + \lambda\cos(k_{II}L)}{\cos(k_{II}L) + \frac{\lambda}{k_{II}}\sin(k_{II}L)} \\
& & &= \lambda\frac{1-\frac{k_{II}}{\lambda}\tan(k_{II}L)}{1+\frac{\lambda}{k_{II}}\tan(k_{II}L)}
\end{aligned}$$

We insert the result again into Eq. 4.44 in order to obtain the solution in region III.

$$\begin{aligned}
\psi_{III}(x) &= A_I\left[\cos(k_{II}L) + \frac{\lambda}{k_{II}}\sin(k_{II}L)\right]\left[\mathrm{e}^{\lambda(x-L)}\frac{-\lambda-D_3}{-\lambda-\lambda} + \mathrm{e}^{-\lambda(x-L)}\frac{\lambda-D_3}{\lambda+\lambda}\right] \\
&= A_I\left[\cos(k_{II}L) + \frac{\lambda}{k_{II}}\sin(k_{II}L)\right]\left[\mathrm{e}^{\lambda(x-L)}\frac{1}{2}\left(1+\frac{D_3}{\lambda}\right) + \mathrm{e}^{-\lambda(x-L)}\frac{1}{2}\left(1-\frac{D_3}{\lambda}\right)\right]
\end{aligned}$$

The boundary condition require that the prefactor of the exponentially increasing term vanishes, which provides us with a condition for D_3

$$\begin{aligned}
D_3 &= -\lambda \\
\lambda\frac{1-\frac{k_{II}}{\lambda}\tan(k_{II}L)}{1+\frac{\lambda}{k_{II}}\tan(k_{II}L)} &= -\lambda \\
1-\frac{k_{II}}{\lambda}\tan(k_{II}L) &= -1-\frac{\lambda}{k_{II}}\tan(k_{II}L) \\
2 &= \left(\frac{k_{II}}{\lambda}-\frac{\lambda}{k_{II}}\right)\tan(k_{II}L) \\
\frac{2k_{II}\lambda}{k_{II}^2-\lambda^2} &= \tan(k_{II}L)
\end{aligned}$$

Now we introduce the expressions for k_{II} and λ

$$\frac{2\sqrt{-E(E-V)}}{2E-V} = \tan(\frac{1}{\hbar}\sqrt{2m(E-V)}) \tag{D.1}$$

This system of equations cannot be solved analytically, but the structure of the solutions can be visualized graphically as shown in Fig. D.2.

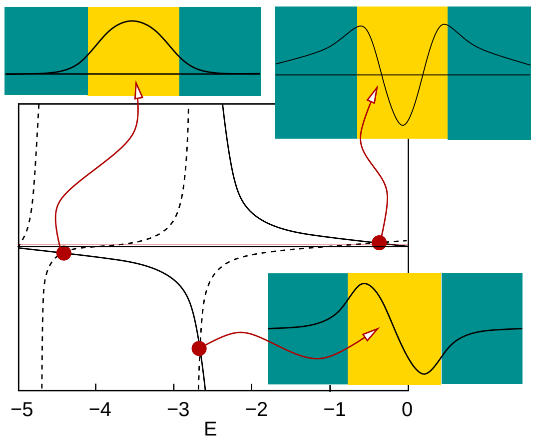


Fig. D.1: Graphical representation of Eq. D.1 and the resulting eigenstates. The left-hand side of Eq. D.1 is the full line, wile the tan-function on the right-hand side is the dashed lines. The crossing points, denoted by the red dots represent the energies of the bound states. The Insets show the corresponding wave functions. The color coding indicates the extent of the well.

Appendix E

Alternative proof of Heisenberg's uncertainty principle

Now let us work out the proof of Heisenberg's uncertainty principle, which is not difficult to follow, but requires a quite clever step. I will show here the proof for the more general uncertainty relation Eq. 7.7, which is not more complicated than the one for the more restricted form Eq. 7.7.

1. Firstly, we show that the pair $\Delta\hat{A} \stackrel{\text{def}}{=} \hat{A} - \langle\hat{A}\rangle$ and $\Delta\hat{B} \stackrel{\text{def}}{=} \hat{B} - \langle\hat{B}\rangle$ obeys the same commutation relations as the operators $\hat{A}$ and $\hat{B}$ themselves

PROOF:

$$\begin{aligned}[\Delta\hat{A}, \Delta\hat{B}] &= [\hat{A} - \langle\hat{A}\rangle, \hat{B} - \langle\hat{B}\rangle] \\ &= (\hat{A}\hat{B} - \hat{A}\langle\hat{B}\rangle - \langle\hat{A}\rangle\hat{B} + \langle\hat{A}\rangle\langle\hat{B}\rangle) - (\hat{B}\hat{A} - \hat{B}\langle\hat{A}\rangle - \langle\hat{B}\rangle\hat{A} + \langle\hat{A}\rangle\langle\hat{B}\rangle) \\ &= [\hat{A}, \hat{B}]\end{aligned}$$

q.e.d

2. Now we show that, for any pair of hermitian operators $\hat{A}$ and $\hat{B}$, the product of expectation values obeys the relation $\sqrt{\langle\psi|\hat{A}^2|\psi\rangle\langle\psi|\hat{Q}^2|\psi\rangle} \geq \frac{1}{2}\Big|\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle\Big|$.

PROOF: We start by defining the operator $\hat{S}(\lambda)=\hat{A}+i\lambda\hat{B}$

$$F(\lambda)\stackrel{\text{def}}{=}\langle\psi|\hat{S}^\dagger\hat{S}|\psi\rangle=\langle\hat{S}\psi|\hat{S}\psi\rangle\geq 0 \tag{E.1}$$

Since $\hat{S}^\dagger=\hat{A}-i\lambda\hat{B}$

$$\begin{aligned}F(\lambda)&=\langle\psi|(\hat{A}-i\lambda\hat{B})(\hat{A}+i\lambda\hat{B})|\psi\rangle\\&=\langle\psi|(\hat{A}^2+i\lambda[\hat{A},\hat{B}]_-+\lambda^2\hat{B}^2)|\psi\rangle\\&=\langle\psi|\hat{A}^2|\psi\rangle+\lambda\langle\psi|i[\hat{A},\hat{B}]|\psi\rangle+\lambda^2\langle\psi|\hat{B}^2|\psi\rangle\stackrel{\text{Eq. }E.1}{\geq}0\end{aligned}$$

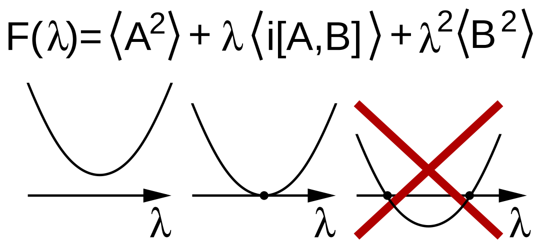
$F(\lambda)$ is a quadratic function in λ. $F(\lambda)$ is real because i times the commutator of two hermitian operators is hermitian. This is shown as follows:

$$\Big(i[\hat{A},\hat{B}]\Big)^\dagger=-i\Big(\hat{A}\hat{B}-\hat{B}\hat{A}\Big)^\dagger=-i\Big(\hat{B}^\dagger\hat{A}^\dagger-\hat{A}^\dagger\hat{B}^\dagger\Big)\stackrel{\hat{A}^\dagger=\hat{A};\hat{B}^\dagger=\hat{B}}{=}-i\Big(\hat{B}\hat{A}-\hat{A}\hat{B}\Big)=i[\hat{A},\hat{B}]$$

Because the expectation value of a hermitian operator is real, also the middle coefficient in $F(\lambda)$ is real.
The parabola is always upwards open, which can be seen by determining the sign of the prefactor of λ^2. We obtain

$$\langle\psi|\hat{B}^2|\psi\rangle=\langle\psi|\hat{B}^2|\psi\rangle\stackrel{\hat{B}=\hat{B}^\dagger}{=}\langle\hat{B}\psi|\hat{B}\psi\rangle\geq 0\qquad\Rightarrow\qquad\frac{d^2F}{d\lambda^2}\geq 0$$

The requirement that the function is always non-negative tells us that the function has at most one **real** zero. (A parabola has always two complex zeros.)



The zeros of $F(\lambda)$ are

$$\lambda_\pm=-\frac{\langle\psi|i[\hat{A},\hat{B}]|\psi\rangle}{2\langle\hat{A}^2\rangle}\pm\sqrt{\left(\frac{\langle\psi|i[\hat{A},\hat{B}]|\psi\rangle}{2\langle\hat{A}^2\rangle}\right)^2-\frac{\langle\hat{B}^2\rangle}{\langle\hat{A}^2\rangle}}$$

The expression under the square-root on the right-hand side must be zero or negative, because otherwise there would be two real zeros of the parabola, which would violate the requirement that the parabola is non-negative for all values of λ.

$$\begin{aligned}&\left(\frac{\langle\psi|i[\hat{A},\hat{B}]|\psi\rangle}{2\langle\psi|\hat{A}^2|\psi\rangle}\right)^2-\frac{\langle\psi|\hat{B}^2|\psi\rangle}{\langle\psi|\hat{A}^2|\psi\rangle}\leq 0\\\Rightarrow\qquad&\sqrt{\langle\psi|\hat{B}^2|\psi\rangle}\sqrt{\langle\psi|\hat{A}^2|\psi\rangle}\geq\frac{1}{2}\Big|\langle\psi|i[\hat{A},\hat{B}]|\psi\rangle\Big|\end{aligned}$$

3. Similarly we can insert $\Delta\hat{A}$ for $\hat{A}$ and $\Delta\hat{B}$ for $\hat{B}$ and obtain Heisenberg's uncertainty principle Eq. 7.7.

$$\sqrt{\langle\psi|(\Delta\hat{A})^2|\psi\rangle}\sqrt{\langle\psi|(\Delta\hat{B})^2|\psi\rangle}\geq\frac{1}{2}\Big|\langle\psi|i[\hat{A},\hat{B}]|\psi\rangle\Big| \tag{E.2}$$

4. The original formulation of Heisenberg's uncertainty principle is obtained as a special case of

Eq. E.2. For a generalized coordinate $\hat{Q}$ and its canonical conjugate momentum $\hat{P}$ we obtain

$$\sqrt{\langle\psi|(\Delta\hat{Q})^2|\psi\rangle}\sqrt{\langle\psi|(\Delta\hat{P})^2|\psi\rangle} \geq \frac{1}{2}\left|\langle\psi|i[\hat{P},\hat{Q}]|\psi\rangle\right| = \frac{\hbar}{2} \tag{E.3}$$

because of their specific commutator relation. *q.e.d.*

Appendix F

Nodal-line theorem (Knotensatz)

Here we present two nodal theorems. One is limited to ordinary differential equations of second order, such as the radial Schrödinger equation for an atom. The other is more general and is applicable also to partial differential equations such as the Schrödinger equation in three dimensions or the one for several particles. The conclusion of the latter nodal theorem is more limited, which is the reason for presenting both.

F.1 Nodal theorem for ordinary differential equations of second order

For the eigenvalue problem of an ordinary second order differential equation with homogeneous boundary conditions[1], the n-th eigenstate has $n-1$ nodes within the considered interval.

The source is the book by Courant and Hilbert *"Die Methoden der Mathematischen Physik I"*, section 6.

F.1.1 Proof

Here we show that the nodes of an inhomogeneous radial Schrödinger equation with a local potential and node-less inhomogeneity only migrate with increasing energy from the outer boundary inward and that two nodes neither annihilate nor does a pair of nodes form spontaneously.

The inhomogeneous Schrödinger equation $(\hat{p}^2/(2m_e)+\hat{V}-E)|\phi(E)\rangle=0$ can be written in the form

$$\left[-\frac{\hbar^2}{2m_e}\partial_r^2+v(r)-E\right]f(E,r)=g(r) \tag{F.1}$$

with

$$\begin{aligned}\phi(E,\vec{r})&=\frac{1}{|\vec{r}|}f(E,|\vec{r}|)Y_{\ell,m}(\vec{r})\\ V(\vec{r})&=v(|\vec{r}|)-\frac{\hbar^2\ell(\ell+1)}{2m_e|\vec{r}|^2}\end{aligned}$$

We impose specific initial conditions for the differential equation, namely that value and radial derivative of $f(E,r)$ at the origin is energy independent.

It is helpful to realize the analogy of the inhomogeneous, radial Schrödinger equation to Newtons equation of motion for a driven harmonic oscillator with a time-dependent spring constant.

[1]The value vanishes at the bounds of the interval

The proof will be based on the study of the motion of nodes with increasing energy. Let us consider the position $\bar{r}(E)$ of a particular node. From $f(E,\bar{r}(E))=0$, we obtain

$$\partial_E\bar{r}=-\dot{f}(E,\bar{r})/\partial_r f(E,\bar{r}) \tag{F.2}$$

where we define the energy derivative of $f(E,r)$ as $\dot{f}(E,r)=\partial f(E,r)/\partial E$. The energy derivative $\dot{f}$ obeys itself an inhomogeneous Schrödinger equation

$$[-\frac{\hbar^2}{2m_e}\partial_r^2+v(r)-E]\dot{f}(E,r)=f(E,r) \tag{F.3}$$

Because of the condition that value and derivative of f are energy independent, value and derivative of its energy derivative vanish. Thus we obtain for a node position $\bar{r}$

$$\begin{aligned}
\dot{f}(E,\bar{r})\partial_r f(E,\bar{r}) &= \left[\dot{f}\partial_r f - f\partial_r\dot{f}\right]_0^{\bar{r}} \\
&= \int_0^{\bar{r}} dr\,\partial_r\left(\dot{f}\partial_r f - f\partial_r\dot{f}\right) \\
&= \int_0^{\bar{r}} dr\,\left(\dot{f}\partial_r^2 f - f\partial_r^2\dot{f}\right) \\
&= \frac{2m_e}{\hbar^2}\int_0^{\bar{r}} dr\,\left[\dot{f}(v-\epsilon)f - f(v-\epsilon)\dot{f}+f^2\right] \\
&= \frac{2m_e}{\hbar^2}\int_0^{\bar{r}} dr\,f^2(r)
\end{aligned}$$

Thus we find

$$\partial_E\bar{r}=-\frac{2m_e}{\hbar^2}\frac{1}{|\partial_r f(E,\bar{r})|^2}\int_0^{\bar{r}} dr\,f^2(r) \tag{F.4}$$

The integral is always positive. This implies in turn, that the nodes move inward with increasing energy, and that the distance between the nodes shrinks with increasing energy.

F.2 Courant's Nodal Line Theorem

Consider a eigenvalue problem of the form

$$(\hat{L}-Ek(x))f(\vec{x})=0 \tag{F.5}$$

with a linear, hermitian differential operator $\hat{L}=L(\vec{\nabla},\vec{x})$, a positive overlap type term $k(\vec{x})>0$, and homogeneous boundary conditions on the region G, that is $f(\vec{x})=0$ on the boundary of the region G.

Courant shows[21][2] [3] that the n-th eigenstate $f_n(x)$ divides the region G into no more than n subregions.

In contrast to the regular nodal line theorem, this theorem is not limited to ordinary differential equations of second order, but is valid for partial differential equations of any order.

[2]"Courant's nodal line theorem" is also covered in the textbook: R. Courant, D. Hilbert, Methoden der Mathematischen Physik Vol. 1.

[3]In Courant's paper [21] appears to be a small typo on p.82 in the bottom line: u_i should be u_n.

Specialization

The proof is carried out for a second-order, ordinary differential equation[4] of the form

$$(-\partial_x^2 + v(x) - E)f(x) = 0 \tag{F.6}$$

The generalization is supposed to be straight forward.

For the time being, I am assuming that all functions are real valued.

One

In a first step, Courant selects a specific band index n and defines a functional F of several functions

$$\begin{aligned} F_n([g_1], \ldots, [g_{n-1}]) := \min_{[\phi(x)],\varepsilon,\{\gamma_j\}} \Biggl\{ \int_a^b dx\ \Bigl[(\partial_x\phi(x))^2 + v(x)\phi^2(x)\Bigr] \\ -\varepsilon\Bigl[\int_a^b dx\ \phi^2(x) - 1\Bigr] - \sum_{j=1}^{n-1}\gamma_j\Bigl[\int_a^b dx\ g_j(x)\phi(x)\Bigr]\Biggr\} \end{aligned} \tag{F.7}$$

The constraints, namely the orthogonality to the $g_j(x)$ and the normalization constraint are enforced by the method of Lagrange multipliers. The constraint conditions require that the derivative with respect to the Lagrange multipliers ε and γ_j vanish.

The equilibrium condition for $\phi(x)$ is

$$\Bigl(-\partial_x^2 + v(x) - \varepsilon\Bigr)\phi(x) = \frac{1}{2}\sum_j \gamma_j g_j(x) \tag{F.8}$$

The energy ε is equal to the value of the functional, when the equilibrium condition is fulfilled. This is shown by multiplication of Eq. F.8 from the left with ϕ and comparison of the resulting expression with Eq. F.7.

Two

In a second step, Courant searches for the maximum of F, which produces the condition

$$\frac{\delta F}{\delta g_j(x)} = \gamma_j\phi(x) = 0 \tag{F.9}$$

which can only be fulfilled, if all Lagrange multipliers γ_j vanish. Thus the corresponding function $\bar{\phi}(x)$ obeys an eigenvalue equation

$$\Bigl(-\partial_x^2 + v(x) - E\Bigr)\bar{\phi}(x) = 0 \tag{F.10}$$

Courant concludes that the resulting functions $g_j(x)$ correspond to the $n-1$ lowest eigenstates, and that $\phi(x)$ corresponds to the n-th eigenstate of the differential equation.

If the functions g_j do not span the space of the $n-1$ lowest eigenstates, the function ϕ could take an admixture of the missing eigenstate and thus reduce its eigenvalue E. As a consequence E would not be the maximum of the functional F.

Incomplete: Show that the ϕ_n is the highest state....

[4] Courant's original proof is more general than what is shown here.

Three

Consider now a subset G' of the interval $G := [a, b]$, and formulate a differential equation of the form Eq. F.6 on this subset.

The analogous functional F' can be expressed by the original functional F by restricting ϕ to vanish outside of G'. Thus for every set of arguments for the functional $F' \geq F$, because the search of functions is restricted to those that disappear outside G'.

Because the functional F' lies for all arguments above F, also the maximum value $E' = \max F'$ for the new functional lies above the maximum value $E = \max F$ for the original problem.

Four

Let m be the number of sub-regions G_j of G that are divided by zeros of the n-th eigenvalue $f_n(x)$ and let us assume that $m > n$ in contradiction to the claim of Courant's theorem.

For each such subregion G_j, we define a function $w_j(x)$, that is zero outside of G_j and that is inside proportional to the n-th eigenstate $f_n(x)$. The proportionality constant is determined by the normalization of $w_j(x)$: Each function $w_j(x)$ is normalized so that $\int_G dx\; w_j^2(x) = 1$.

The normalization condition, together with the differential equation, shows that

$$\int_a^b dx\; \Big[(\partial_x w_j(x))^2 + v(x) w_j^2(x)\Big] = E_n \tag{F.11}$$

where E_n is the eigenvalue corresponding to the n-th eigenstate $f_n(x)$.

We can now define a function

$$W_{\vec{c}}(x) = \sum_j w_j(x) c_j \tag{F.12}$$

The normalization condition for $W_{\vec{c}}(x)$ can be fulfilled by choosing $\sum_j c_j^2 = 1$. With this condition we find

$$\begin{aligned}\int_a^b dx\; \Big[(\partial_x W_{\vec{c}}(x))^2 + v(x) W_{\vec{c}}^2(x)\Big] &= \sum_j c_j^2 \int_a^b dx\; \Big[(\partial_x w_j(x))^2 + v(x) w_j^2(x)\Big] \\ &= \sum_j c_j^2 E = E \end{aligned} \tag{F.13}$$

Let us fix the arguments $[g_1, \ldots, g_{n-1}]$ of the function F. The orthogonality condition $W_{\vec{c}}(x) g_j = 0$ determines, together with normalization constraint n components of $\vec{c}$.

If the number m of regions G_j, that are separated by the zeros of $f_n(x)$, one can construct $m!/(n!(m-n)!)$ different vectors $\vec{c}$, that satisfy all the n constraints. Each of these solutions is constructed from selecting n out of the m regions, and by setting the components of $\vec{c}$ in the remaining $m - n$ regions to zero. The vector $\vec{c}$ is determined by (1) setting the components in the remaining regions to zero, by (2) imposing $n - 1$ orthogonality constraints and by (3) imposing the normalization constraint. Because the number of distinct selections of n regions out of m is $m!/(n!(m-n)!$, this is the number of distinct, that is linear independent solutions

$$b_j(x) = W_{\vec{c}^{(j)}}(x) \qquad \text{for } j = 1, \ldots, \frac{m!}{n!(m-n)!} \tag{F.14}$$

corresponding to the vectors $\vec{c}^{(1)}, \vec{c}^{(2)}, \ldots$. All these functions have the same value E_n of the functional $F_n([g_1], \ldots)$ as the state $f_n(x)$.

Five

Now we collect an eigenstate with the next higher eigenvalue E_p, with $E_p > E_n$. **We assume that $p \leq m!/(n!(m-n)!)$.**

Now we turn to the functional $F_p([g_1], \ldots, [g_{p-1}])$ for the p-th eigenstate. Out of the functions $b_j(x)$ we can form a superposition

$$\phi(x) = \sum_{j=1}^{m!/(n!(m-n)!)} b_j(x) k_j \tag{F.15}$$

so that the orthogonality to the $p-1$ arguments $g_j(x)$ is fulfilled and the normalization condition is obeyed. As there are p conditions, the ansatz Eq. F.15 must have at least p terms.

Note, that any set of arguments, the value of the functional F_p is smaller than E_n, because we just have constructed one solution $\phi(x)$, which obeys all constraints and for which the integral is equal to E_n. Because the value of the functional F_p is the minimum for all functions Φ that obey the constraints, we can deduce that

$$E_p = \max_{[g_1], \ldots, [g_{p-1}]} F_p([g_1], \ldots, [g_p]) \leq E_n \tag{F.16}$$

This however is in contradiction to the requirement that $E_p > E_n$.

Conclusion

To avoid the contradiction, the number of areas G_j that are separated by the node planes of $f_n(x)$, must be smaller or equal n.

Addition Peter Blöchl: In case of degeneracy we would require that the number of areas must have a value m, so that $m!/(n!(n-m)!) < p$, where E_p is the next larger eigenvalue than E_n.

Appendix G

Spherical harmonics

G.1 Spherical harmonics addition theorem

$$P_\ell(\frac{\vec{r}_1}{|\vec{r}_1|}\frac{\vec{r}_2}{|\vec{r}_2|}) = \frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell} Y^*_{\ell,m}(\vec{r}_1)Y_{\ell,m}(\vec{r}_2)$$

where $P_\ell(x)$ are the Legendre polynomial as defined in Eq. 11.26 on p. 176.

Proof: Rotate the coordinate system so that $\vec{r}_1$ points in z-direction.

$$\begin{aligned}
P_\ell(\frac{r_2\vec{e}_z}{|\vec{r}_2|}) &= \frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell} Y^*_{\ell,m}(\vec{e}_z)Y_{\ell,m}(\vec{r}_2) \\
&= \frac{4\pi}{2\ell+1}]Y^*_{\ell,0}(\vec{e}_z)Y_{\ell,0}(\vec{r}_2) \\
&\stackrel{\text{Eq. 11.23}}{=} \frac{4\pi}{2\ell+1}\frac{2\ell+1}{4\pi}P^0_\ell(1)P^0_\ell(\cos(\theta)) \\
&\stackrel{\text{Eq. 11.25}}{=} P_\ell(1)P_\ell(\cos(\theta)) \\
&\stackrel{\text{Eq. 11.26}}{=} \frac{1}{2^\ell\ell!}\left[\partial^\ell_u(u^2-1)^\ell\right]_{u=1}P_\ell(\cos(\theta))
\end{aligned}$$

To complete the proof, we need to show that $P_\ell(1) = 1$, which is not done yet....

G.2 Unsöld's Theorem

[49]

$$\frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell} Y^*_{\ell,m}(\vec{r})Y_{\ell,m}(\vec{r}) = 1$$

G.3 Condon-Shortley phase

The Condon-Shortley phase $(-1)^m$ is often introduced in quantum mechanical calculations. It can be introduced either the in the definition of the associated Legendre polynomials or in the expression of the spherical harmonics expressed by associated Legendre polynomials.

Appendix H

Time-inversion symmetry

Time inversion symmetry says that it is not possible from a conservative classical trajectory to find out if it time is running forward or backward in time. If we take a physical trajectory $\vec{x}(t)$ and let the time run backwards, i.e. $\vec{x'}(t) = \vec{x}(-t)$, the new trajectory $\vec{x'}(t)$ still fulfills the equations of motion.

Electrodynamics and gravitation obey time-inversion symmetry exactly. However, while time time-inversion symmetry is one of the most fundamental properties of natural laws, it is, taken alone, not a fundamental symmetry of nature: The weak interaction, which is, for example, responsible for the β decay of nuclei, violates it. Time inversion must be replaced by the weaker CPT-inversion symmetry. This is the so-called **CPT-theorem** posed by Gerhart Lüders and Wolfgang Pauli. The CPT theorem says that the fundamental laws of nature must obey a symmetry under simultaneous application of three operations:

- charge inversion (C)
- space inversion (P for Parity)
- time inversion (T)

The CPT theorem is based on the assumptions of Lorentz invariance, causality, locality and the existence of a Hamilton operator that is bounded by below. Electrodynamics and gravitation are symmetric under the three symmetry operations individually.

As we are not concerned with weak interactions we can assume exact time inversion symmetry.

H.1 Schrödinger equation

Let us now investigate what time inversion symmetry implies in quantum mechanics:

Let us consider the Schrödinger equation in a magnetic field

$$i\hbar\partial_t\Psi(\vec{r},t) = \left[\frac{(\frac{\hbar}{i}\vec{\nabla} - q\vec{A})^2}{2m} + q\Phi\right]\Psi(\vec{r},t) \tag{H.1}$$

Let us take the complex conjugate of the Eq. H.1

$$-i\hbar\partial_t\Psi^*(\vec{r},t) = \left[\frac{(-\frac{\hbar}{i}\vec{\nabla} - q\vec{A})^2}{2m} + q\Phi\right]\Psi^*(\vec{r},t) = \left[\frac{(\frac{\hbar}{i}\vec{\nabla} + q\vec{A})^2}{2m} + q\Phi\right]\Psi^*(\vec{r},t) \tag{H.2}$$

Next we look for the equation obeyed by $\psi(\vec{r},-t)$, if Eq. H.2 holds

$$i\hbar\partial_t\Psi^*(\vec{r},-t) = \left[\frac{(\frac{\hbar}{i}\vec{\nabla} + q\vec{A}(\vec{r},-t))^2}{2m} + q\Phi(\vec{r},-t)\right]\Psi^*(\vec{r},-t) \tag{H.3}$$

One can see immediately that Eq. H.3 is identical to Eq. H.1, when the time and simultaneously the vector potential $\vec{A}$ are reverted. Thus, the Schrödinger equation is symmetric under the time inversion symmetry as stated below:

TIME-INVERSION

$$\vec{A}(\vec{r}, t) \rightarrow -\vec{A}(\vec{r}, -t) \tag{H.4}$$

$$\Phi(\vec{r}, t) \rightarrow \Phi(\vec{r}, -t) \tag{H.5}$$

$$\Psi(\vec{r}, t) \rightarrow \Psi^*(\vec{r}, -t) \tag{H.6}$$

For the time-independent Schrödinger equation we obtain

$$\Psi(\vec{r}, t) = \Psi_\epsilon(\vec{r})\mathrm{e}^{-\frac{i}{\hbar}\epsilon t} \tag{H.7}$$

so that

$$\Psi^*(\vec{r}, -t) \stackrel{\text{Eq. } H.7}{=} \Psi^*_\epsilon(\vec{r})\mathrm{e}^{-\frac{i}{\hbar}\epsilon t} \tag{H.8}$$

Thus, the time-inversion symmetry applied to energy eigenstates has the effect that the wave function is turned into its complex conjugate.

Let us look at the problem from stationary Schrödinger equation.

$$\left[\frac{(\frac{\hbar}{i}\vec{\nabla} - q\vec{A})^2}{2m} + q\Phi - \epsilon\right]\Psi_\epsilon(\vec{r}) = 0 \tag{H.9}$$

We take the complex conjugate of this equation

$$\left[\frac{(\frac{\hbar}{i}\vec{\nabla} + q\vec{A})^2}{2m} + q\Phi - \epsilon\right]\Psi^*_\epsilon(\vec{r}) = 0 \tag{H.10}$$

We observe that the complex conjugate of the wave function solves the same Schrödinger equation with the magnetic field reversed.

This implies that, in the absence of a vector potential $\vec{A}$, the complex conjugate is also a solution of the original wave function. When these two degenerate solutions are superimposed, one obtains another solution of the same Schrödinger equation. Hence, also the real part and the imaginary part are solutions of the Schrödinger equation. Thus, in the absence of a magnetic field, the wave functions can be assumed to be purely real.

REAL WAVE FUNCTIONS

In the absence of magnetic fields, the wave function of the non-relativistic Schrödinger equation can be chosen as real functions.

H.2 Pauli equation

The proper theory of electrons is the **Dirac equation**, which describes electrons by a four component spinor, that describes spin up and spin-down electrons as well as their antiparticles, the positrons. In the non-relativistic limit, electrons and positrons become independent. In this limit electrons and positrons obey the so-called Pauli equation.

The **Pauli equation** has the form

$$i\hbar\partial_t|\psi(t)\rangle = \left[\frac{(\vec{p}-q\vec{A})^2}{2m_e} + q\Phi - \frac{q}{m_e}\hat{\vec{S}}\vec{B}\right]|\psi\rangle \tag{H.11}$$

The wave function is now a two-component spinor with a spin-up and a spin-down component. The spin-operator is represented by the Pauli matrices

$$\vec{S} \stackrel{\text{Eq. ??}}{=} \frac{\hbar}{2}\left(\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\right)$$

The Pauli matrices are given in Eq. 6.23 on 112. The magnetic field is related to the vector potential via $\vec{B} = \vec{\nabla} \times \vec{A}$.

Expressed with explicit spinor components, the Pauli equation has the form

$$i\hbar\partial_t\psi(\vec{r},\sigma,t) = \sum_{\sigma'}\left[\underbrace{\left(\frac{(\frac{\hbar}{i}\vec{\nabla}-q\vec{A})^2}{2m_e} + q\Phi\right)\delta_{\sigma,\sigma'}}_{\hat{H}_0} - \frac{q}{m_e}\vec{B}\vec{S}_{\sigma,\sigma'}\right]\psi(\vec{r},\sigma',t) \tag{H.12}$$

We proceed as we did for the Schrödinger equation by taking the complex conjugate of the Pauli equation Eq. H.12

$$-i\hbar\partial_t\psi^*(\vec{r},\sigma,t) = \sum_{\sigma}\left[\left(\frac{(-\frac{\hbar}{i}\vec{\nabla}-q\vec{A})^2}{2m_e} + q\Phi\right)\delta_{\sigma,\sigma'} - \frac{q}{m_e}\vec{B}\vec{S}^*_{\sigma,\sigma'}\right]\psi^*(\vec{r},\sigma',t) \tag{H.13}$$

Now we revert the time argument[1] in Eq. H.13 and perform the transformation of the potentials $\vec{A}'(\vec{r},t) = -\vec{A}(\vec{r},-t)$ and $\Phi'(\vec{r},t) = \Phi(\vec{r},-t)$. The transformed magnetic field is $\vec{B}'(\vec{r},t) = \vec{\nabla} \times \vec{A}'(\vec{r},t) = -\vec{\nabla} \times \vec{A}(\vec{r},-t) = -\vec{B}(\vec{r},-t)$.

$$i\hbar\partial_t\psi^*(\vec{r},\sigma,-t) = \sum_{\sigma'}\left[\underbrace{\left(\frac{(\frac{\hbar}{i}\vec{\nabla}-q\vec{A}'(\vec{r},t))^2}{2m_e} + q\Phi'(\vec{r},t)\right)\delta_{\sigma,\sigma'}}_{\hat{H}'_0} + \frac{q}{m_e}\vec{B}'(\vec{r},t)\vec{S}^*_{\sigma,\sigma'}\right]\psi^*(\vec{r},\sigma',-t) \tag{H.14}$$

We observe, that it is no more sufficient to replace the wave function by its complex conjugate of the time-reverted function as in the Schrödinger equation. This is because the complex conjugate of the spin operator $\vec{S}_{\sigma,\sigma'}$ is not identical to its complex conjugate. The y-component is purely imaginary, because the corresponding Pauli matrix is purely imaginary. Thus, the y-component changes its sign, when the complex conjugate is taken.

In order to find the transformation of the wave functions, let us rewrite the equation in components. The original equation Eq. H.12 written in components looks like

$$(i\hbar\partial_t - \hat{H}_0)\begin{pmatrix}\psi(\vec{r},\uparrow,t)\\ \psi(\vec{r},\downarrow,t)\end{pmatrix} \stackrel{\text{Eq. }H.12}{=} -\frac{\hbar q}{2m_e}\begin{pmatrix}B_z & B_x - iB_y\\ B_x + iB_y & -B_z\end{pmatrix}\begin{pmatrix}\psi(\vec{r},\uparrow,t)\\ \psi(\vec{r},\downarrow,t)\end{pmatrix} \tag{H.15}$$

The last equation, Eq. H.14, written in component notation has the form

$$(i\hbar\partial_t - \hat{H}'_0)\begin{pmatrix}\psi^*(\vec{r},\uparrow,-t)\\ \psi^*(\vec{r},\downarrow,-t)\end{pmatrix} \stackrel{\text{Eq. }H.14}{=} \underbrace{-\frac{\hbar q}{2m_e}\begin{pmatrix}-B'_z & -B'_x - iB'_y\\ -B'_x + iB'_y & B'_z\end{pmatrix}}_{+\frac{q}{m_e}\vec{B}'\vec{S}^*}\begin{pmatrix}\psi^*(\vec{r},\uparrow,-t)\\ \psi^*(\vec{r},\downarrow,-t)\end{pmatrix} \tag{H.16}$$

[1] Consider a replacement $t = -t'$. Thus, $\partial_{t'} = -\partial_t$ and $\psi(t) = \psi(-t')$. After the transformation we drop the prime.

In order to get an idea on how to proceed, let us consider the special case $B'_x = B'_y = 0$. In this case the diagonal elements of the equation could be brought into the form of Eq. H.15 by interchanging the spin indices of the wave functions.

Guided by this idea, we interchange the spin indices of Eq. H.16. This interchanges the columns and the rows of the 2×2 matrix. We obtain

$$\left(i\hbar\partial_t - \hat{H}'\right)\begin{pmatrix}\psi^*(\vec{r},\downarrow,-t)\\ \psi^*(\vec{r},\uparrow,-t)\end{pmatrix} = -\frac{\hbar q}{2m_e}\begin{pmatrix}B'_z & -B'_x + iB'_y\\ -B'_x - iB'_y & -B'_z\end{pmatrix}\begin{pmatrix}\psi^*(\vec{r},\downarrow,-t)\\ \psi^*(\vec{r},\uparrow,-t)\end{pmatrix}$$

The off-diagonal elements of the equation still differ by the sign from those in Eq. H.15. This problem can be remedied by replacing the upper spinor component by its negative. While rewriting the equation for the new convention, the sign of the off-diagonal elements of the matrix is changed.

$$\left(i\hbar\partial_t - \hat{H}'\right)\begin{pmatrix}-\psi^*(\vec{r},\downarrow,-t)\\ \psi^*(\vec{r},\uparrow,-t)\end{pmatrix} = -\frac{\hbar q}{2m_e}\begin{pmatrix}B'_z & B'_x - iB'_y\\ B'_x + iB'_y & -B'_z\end{pmatrix}\begin{pmatrix}-\psi^*(\vec{r},\downarrow,-t)\\ \psi^*(\vec{r},\uparrow,-t)\end{pmatrix} \quad \text{(H.17)}$$

Thus, we arrived at the desired form Eq. H.15.

By comparing Eq. H.17 with the original Pauli equation Eq. H.15, we see that the for any solution of the Pauli equation also the result of the time inversion operation Eq. H.18 for two-component spinors and electromagnetic fields is a solution of the same Pauli equation.

TIME-INVERSION FOR TWO-COMPONENT SPINORS AND ELECTROMAGNETIC FIELDS

$$\begin{aligned}\psi'(\vec{r},\uparrow,t) &= -\psi^*(\vec{r},\downarrow,-t)\\ \psi'(\vec{r},\downarrow,t) &= +\psi^*(\vec{r},\uparrow,-t)\\ \vec{A}'(\vec{r},t) &= -\vec{A}(\vec{r},-t)\\ \Phi'(\vec{r},t) &= \Phi(\vec{r},-t)\end{aligned} \quad \text{(H.18)}$$

If we investigate the resulting transformation of the spin-expectation values, we see that the spin is inverted. This is expected under time-inversion symmetry, if we interpret the spin as a angular momentum. If we invert the time, the particle is spinning in the opposite direction.

H.3 Time inversion for Bloch states

Often we require the implications of this symmetry for wave in a representation of Bloch waves. A Bloch state is given by as product of a periodic function $u_{\vec{k},n}$ and a phase factor $e^{i\vec{k}\vec{r}}$, so that

$$\psi_{\vec{k},n}(\vec{r},\sigma,t) = u_{\vec{k},n}(\vec{r},\sigma)e^{i(\vec{k}\vec{r}-\vec{\omega}_n t)} \quad \text{(H.19)}$$

If time inversion symmetry is obeyed, for example in the absence of magnetic fields with static potentials, we can use $\psi(\vec{r},\uparrow,t) = -\psi^*(\vec{r},\downarrow,-t)$ and $\psi(\vec{r},\downarrow,t) = -\psi^*(\vec{r},\uparrow,-t)$, so that

$$\begin{aligned}u_{\vec{k},n}(\vec{r},\uparrow)e^{i(\vec{k}\vec{r}-\vec{\omega}t)} = \psi(\vec{r},\uparrow,t) = -\psi^*(\vec{r},\downarrow,-t) &= -u^*_{-\vec{k},n}(\vec{r},\downarrow)e^{i(\vec{k}\vec{r}-\vec{\omega}t)}\\ u_{\vec{k},n}(\vec{r},\uparrow) &= -u^*_{-\vec{k},n}(\vec{r},\downarrow)\end{aligned} \quad \text{(H.20)}$$

and

$$\begin{aligned}u_{\vec{k},n}(\vec{r},\downarrow)e^{i(\vec{k}\vec{r}-\vec{\omega}t)} = \psi(\vec{r},\downarrow,t) = \psi^*(\vec{r},\uparrow,-t) &= u^*_{-\vec{k},n}(\vec{r},\uparrow)e^{i(\vec{k}\vec{r}-\vec{\omega}t)}\\ u_{\vec{k},n}(\vec{r},\downarrow) &= u^*_{-\vec{k},n}(\vec{r},\uparrow)\end{aligned} \quad \text{(H.21)}$$

Thus, we obtain for the periodic parts of the wave functions

$$\begin{pmatrix} u_{\vec{k},n}(\vec{r},\uparrow) \\ u_{\vec{k},n}(\vec{r},\downarrow) \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_{-\vec{k},n}(\vec{r},\uparrow) \\ u_{-\vec{k},n}(\vec{r},\downarrow) \end{pmatrix}^* \tag{H.22}$$

This is an expression of **Kramers' theorem**[50][2]: For static potentials and in the absence of magnetic fields, the energy levels in a crystal obey

$$\epsilon_n(\vec{k}) = \epsilon_n(-\vec{k}) \tag{H.23}$$

[2]Original paper: Kramers, H. A. (1930). "Théorie générale de la rotation paramagnétique dans les cristaux". Proceedings of the Royal Netherlands Academy of Arts and Sciences (in French). 33 (6–10): 959-972.

Appendix I

Rate equation and beyond time-dependent perturbation theory

I.1 Rate equation

I.1.1 Approximate rate equation from Fermi's Golden rule

In order to obtain the time evolution of the occupations under a constant perturbation $\hat{W}$, we may divide the perturbation into small time slices and evaluate the change after each time slice from Eq. **??**. The time slices must still be much longer than the wave length of the oscillation, so that we can make the assumption of mono-chromatic pulses. If the time slices have the duration Δ, the envelope function for that region would have the value $\frac{1}{\Delta}$ and the perturbation has the value $\hat{W}_0^\Delta = \Delta \cdot \hat{W}_0$, where $\hat{W}_0^\Delta$ is the perturbation without envelope function as it occurs in Eq. **??** for a given time slice, while $\hat{W}_0 e^{i\omega t}$ is the physical perturbation.

One obtains the probabilities

$$P_m(t+\Delta) = P_m(t) + \sum_n \Big(\Lambda_{m,n} P_n(t) - \Lambda_{n,m} P_m(t) \Big)$$

where

$$\Lambda_{m,n} = \Delta \frac{2\pi}{\hbar} \left| \langle \phi_m | \hat{W}_0 | \phi_n \rangle \right|^2 \delta(E_n - E_m - \hbar\omega)$$

Thus we obtain

RATE EQUATION USING FERMI'S GOLDEN RULE

$$\partial_t P_m(t) = \sum_n \Big(\Gamma_{m,n} P_n(t) - \Gamma_{n,m} P_m(t) \Big)$$

with

$$\Gamma_{m,n} = \frac{2\pi}{\hbar} \left| \langle \phi_m | \hat{W} | \phi_n \rangle \right|^2 \delta(E_n - E_m - \hbar\omega)$$

Note that this approximation rests on a number of seemingly contradicting assumptions.

I.1.2 Rate equation in the random-phase approximation

Whereas the derivation of section I.1.1 suffers from many approximations that are difficult to control, it is also possible to obtain a rate equation only by imposing the random phase approximation Eq. **??** and Eq. **??**.

A rate equation can also be obtained by comparing the transition matrix element with the result of the kinetic equation Eq. **??**.

$$\partial_t P_m(t) \stackrel{\text{Eq. ??}}{=} \sum_n \Big(\Gamma_{m,n} P_n(t) - \Gamma_{n,m} P_m(t)\Big) = \sum_n \Big(\Gamma_{m,n} - \delta_{m,n} \sum_k \Gamma_{k,m}\Big) P_n(t)$$

Integrating the latter, leads us to an expression for the probabilities at time t

$$P_m(t) = \sum_n \left(\mathrm{e}^{\bar{\boldsymbol{\Gamma}} t}\right)_{m,n} P_n(0)$$

where

$$\bar{\Gamma}_{m,n} = \Gamma_{m,n} - \sum_k \Gamma_{k,m} \delta_{m,n} \,. \tag{I.1}$$

By comparing this expression with Eq. **??** with $\boldsymbol{\Lambda}$ obtained from Eq. **??**, we obtain the identity

$$\mathrm{e}^{\bar{\boldsymbol{\Gamma}} t} = \bar{\Lambda}_{m,n} \stackrel{\text{def}}{=} \Lambda_{m,n}(t) - \sum_k \Lambda_{k,m}(t) \delta_{m,n}$$

$$\bar{\boldsymbol{\Gamma}} = \lim_{t \to \infty} \frac{1}{t} \ln \Big[\bar{\Lambda}(t)\Big]$$

Now that we have an expression for $\bar{\boldsymbol{\Gamma}}$ we need to determine $\boldsymbol{\Gamma}$ from Eq. I.1. It turns out that this equation leaves the diagonal elements of $\boldsymbol{\Gamma}$ undetermined, so that we set them to zero. The rationale is that a transition probability from a state to itself is meaningless. Hence we obtain

$$\Gamma_{m,n} = \begin{cases} \lim_{t \to \infty} \frac{1}{t} \ln \Big[\bar{\Lambda}(t)\Big] & \text{for } m \neq n \\ 0 & \text{for } m = n \end{cases}$$

RATE EQUATION

Within the random phase approximation, the occupations $P_n(t)$ of the individual quantum states $|\phi_m\rangle$ change under a perturbation $\hat{W}_0 \mathrm{e}^{i\omega t}$ according to the rate equation

$$\partial_t P_m(t) = \sum_{n \neq m} \Big(\bar{\Gamma}_{m,n} P_n(t) - \bar{\Gamma}_{n,m} P_m(t)\Big)$$

with

$$\begin{aligned} \bar{\boldsymbol{\Gamma}} &= \lim_{t \to \infty} \frac{1}{t} \ln \Big[\bar{\Lambda}(t)\Big] \\ \bar{\Lambda}_{m,n} &= \Lambda_{m,n} - \delta_{m,n} \sum_k \Lambda_{k,n} \\ \Lambda_{m,n}(t) &\stackrel{\text{Eq. ??}}{=} \left| \frac{1}{\hbar} \int_0^t dt' \langle \phi_m | \hat{W}(t') | \phi_n \rangle \mathrm{e}^{-\frac{i}{\hbar}(E_n - E_m)t'} \right|^2 + O(\hat{W}^3) \end{aligned}$$

The limit $t \to \infty$ assumes that there is a finite decoherence time, $t_{coherence}$ beyond which $\bar{\boldsymbol{\Lambda}}(t)$ becomes exponential. Thus we can evaluate the limiting value for $t > t_{coherence}$.

I.2 Beyond perturbation theory

In order to explore the behavior beyond second order perturbation theory, I performed numerical calculations of a two-state system with energies E_1 and E_2, which is driven from the ground state to the excited state.

For weak perturbations, we see that the energy is pumped into the system with each oscillation as shown in Fig. I.1. The probability after the pulse, rises with the square of the transition matrix element as indicated by the red square line on the right graph of .

For stronger perturbation, we see that the transition probability reduces again. This is the onset of so-called **Rabi-oscillations**, which are seen more clearly in Fig. I.2. On the graph of the right, we see the oscillations in the final probability.

Depending on how strongly the exciting field is detuned from the dipole oscillations with frequency $\omega = (E_2 - E_1)/\hbar$, the Rabi oscillations rises to a finite value or, for an exact match, to one.

For very large perturbations, as seen on the right of Fig. I.2 another phenomenon kicks in, which I attribute to the fact that Rabi-oscillations are of the same frequency as the driving field.

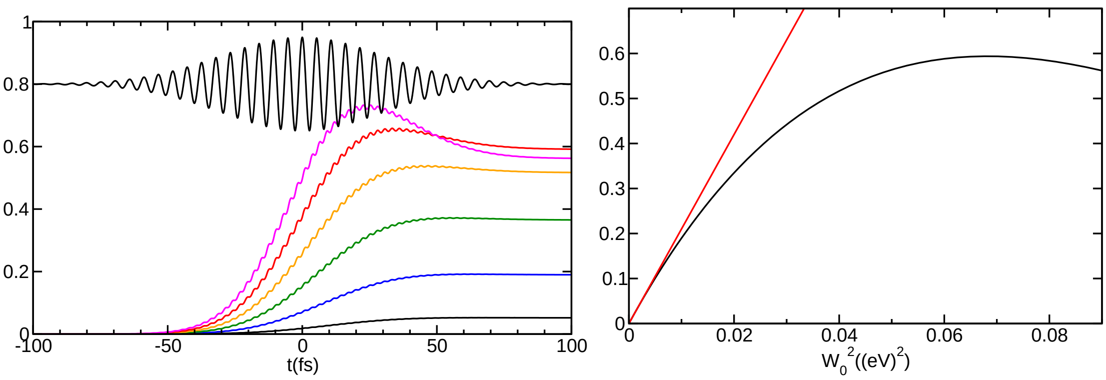


Fig. I.1: Left: Probability to be in the excited state as function of time for different amplitudes of the 50 fs pulse. $E_2 - E_1 = 1$ eV and $\hbar\omega = E_2 - E_1 + 0.02$ eV. The amplitudes are $W_0 = 0.05, 0.10, 0.15, 0.20, 0.25, 0.3$ eV. The exciting wave packet (black) ist scaled arbitrary and displaced vertically. Right: probabilty to be in the excited state after the pulse as function of the squared amplitude W_0^2 of the perturbing field. The red line is a straight line with the initial slope as a guide to the eye.

I.2.1 Code

Editor: The source code is ok. The derivation is in progress.

Diagonalize a (2×2) matrix:

$$\Big(\boldsymbol{h} - \mathbf{1}\epsilon_n\Big)\vec{c}_n = 0 \tag{I.2}$$

Determine eigenvalues

$$\det[\boldsymbol{h} - \mathbf{1}\epsilon_n] = 0 \tag{I.3}$$
$$\Rightarrow \quad (h_{11} - \epsilon)(h_{22} - \epsilon) - h_{12}h_{21} = 0 \tag{I.4}$$
$$\Rightarrow \quad \epsilon^2 - \epsilon(h_{11} + h_{22}) + h_{11}h_{22} - h_{12}h_{21} = 0 \tag{I.5}$$
$$\Rightarrow \quad \epsilon^2 - \mathrm{Tr}[\boldsymbol{h}] + \det[\boldsymbol{h}] = 0 \tag{I.6}$$

$$\epsilon_\pm = \left(\frac{1}{2}\mathrm{Tr}[\boldsymbol{h}]\right) \pm \sqrt{\left(\frac{1}{2}\mathrm{Tr}[\boldsymbol{h}]\right)^2 - \det[\boldsymbol{h}]} \tag{I.7}$$

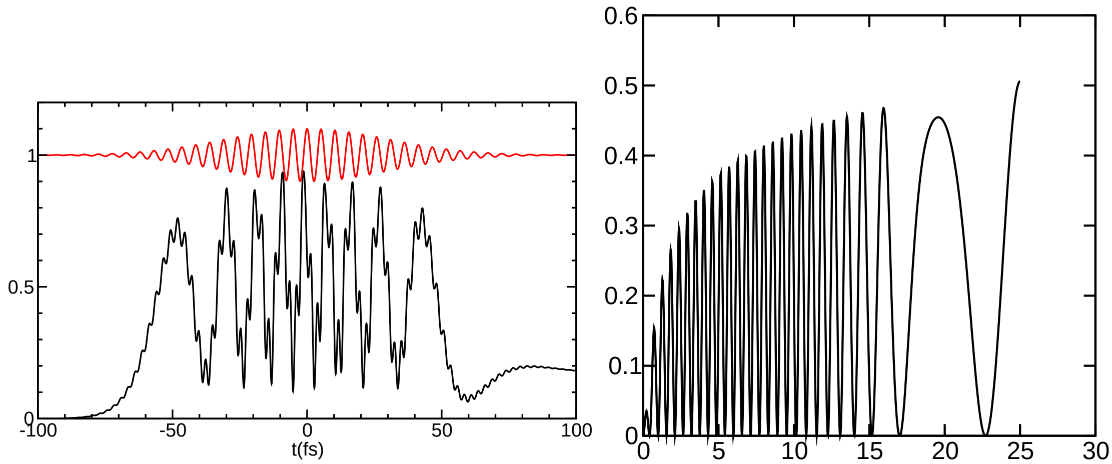


Fig. I.2: Left: Probability to be in the excited state as function of time for a 50 fs pulse with $W_0 = 25$ eV. The level spacing is $E_2 - E_1 = 1$ eV and the light pulse has a frequency with $\hbar\omega = E_2 - E_1 + 0.05$ eV. The exciting wave packet (red) ist scaled arbitrary and shifted vertically. Right: probability to be in the excited state after the pulse as function of the amplitude W_0 of the perturbing field.

Determine eigenvectors: Let us represent the Hamiltonian of the two-state system in terms of unity and the three Pauli matrices defined in Eq. 6.23 on p. 112. I start with some preparation, namely expressing trace and determinant of $\boldsymbol{h}$ in terms of the vector $\vec{a}$

$$\boldsymbol{h} = a_0\mathbf{1} + a_x\boldsymbol{\sigma}_x + a_y\boldsymbol{\sigma}_y + a_z\boldsymbol{\sigma}_z = a_0\mathbf{1} + \vec{a}\vec{\boldsymbol{\sigma}}$$
$$\text{with}\quad a_0 = \frac{1}{2}\mathrm{Tr}[\boldsymbol{h}] \quad\text{and}\quad a_x = \frac{1}{2}(h_{12} + h_{21}) \quad\text{and}\quad a_y = \frac{1}{2i}[h_{21} - h_{12}] \quad\text{and}\quad a_z = \frac{1}{2}[h_{11} - h_{22}] \tag{I.8}$$

$$a_x^2 + a_y^2 + a_z^2 = \underbrace{h_{12}h_{21}}_{a_x^2+a_y^2} + \underbrace{\frac{1}{4}\left(h_{11}^2 + h_{22}^2\right) - \frac{1}{2}h_{11}h_{22}}_{a_z^2}$$
$$= \left(\frac{h_{11} + h_{22}}{2}\right)^2 - \det[\boldsymbol{h}]$$
$$= \left(\frac{1}{2}\mathrm{Tr}[\boldsymbol{h}]\right)^2 - \det[\boldsymbol{h}]$$
$$\det[\boldsymbol{h}] = a_0^2 - \vec{a}^2 \tag{I.9}$$
$$\mathrm{Tr}[\boldsymbol{h}] = 2a_0 \tag{I.10}$$

The eigenvalues of the Hamiltonian obtained above are, expressed in terms of $a_0, \vec{a}$

$$\epsilon_\pm = a_0 \pm \sqrt{\vec{a}^2} = a_0 \pm |\vec{a}| \tag{I.11}$$

To obtain the eigenvectors $\vec{c}_\pm$, we need to solve

$$0 \stackrel{!}{=} \left(\boldsymbol{h} - \epsilon_\pm\mathbf{1}\right)\vec{c}_\pm = \Big[\underbrace{\mathbf{1}a_0 + \vec{a}\vec{\boldsymbol{\sigma}}}_{\boldsymbol{h}} - \mathbf{1}\underbrace{\left(a_0 \pm |\vec{a}|\right)}_{\epsilon_\pm}\Big]\vec{c}_\pm$$
$$\Rightarrow\quad \left[\frac{\vec{a}}{|\vec{a}|}\vec{\boldsymbol{\sigma}} \mp \mathbf{1}\right]\vec{c}_\pm = 0 \tag{I.12}$$

The eigenvectors depend only on the direction of $\vec{a}$, but not on the length of $\vec{a}$ nor on a_0. The eigenvalue equation can therefore be expressed in terms of a unit vector $\vec{e}=\vec{a}/|\vec{a}|$.

Consider the trace-less matrix $\vec{e}\vec{\boldsymbol{\sigma}}$ with $\vec{e}^2=1$

$$0=\begin{pmatrix} e_z\mp 1 & e_x-ie_y \\ e_x+ie_y & -e_z\mp 1\end{pmatrix}\begin{pmatrix} c_{1,\pm}\\ c_{2,\pm}\end{pmatrix}$$

$$\Rightarrow\quad \begin{pmatrix} c_{1,\pm}\\ c_{2,\pm}\end{pmatrix}=\begin{pmatrix} e_x-ie_y\\ -e_z\pm 1\end{pmatrix}\frac{1}{\sqrt{2(1\mp e_z)}}\quad\text{or alternatively}\quad\Rightarrow\quad\begin{pmatrix} c_{1,\pm}\\ c_{2,\pm}\end{pmatrix}=\begin{pmatrix} e_z\pm 1\\ e_x+ie_y\end{pmatrix}\frac{1}{\sqrt{2(1\pm e_z)}}$$

$$\boldsymbol{U}=\begin{pmatrix} e_z+1 & e_x-ie_y\\ e_x+ie_y & -e_z-1\end{pmatrix}\frac{1}{\sqrt{2(1+e_z)}}=\frac{e_x\boldsymbol{\sigma_x}+e_y\boldsymbol{\sigma_y}+(e_z+1)\boldsymbol{\sigma_z}}{\sqrt{2(1+e_z)}}\tag{I.13}$$

With Eq. I.9 and Eq. I.10, one obtains the identity

$$e_x\boldsymbol{\sigma_x}+e_y\boldsymbol{\sigma_y}+e_z\boldsymbol{\sigma_z}=\frac{\boldsymbol{h}-\mathbf{1}\mathrm{Tr}[\boldsymbol{h}]}{\sqrt{\left(\frac{1}{2}\mathrm{Tr}[\boldsymbol{h}]\right)^2-\det[\boldsymbol{h}]}}\tag{I.14}$$

With this identity, the unitary matrix $\boldsymbol{U}$ holding the eigenvectors can be rewritten in the form

$$\boldsymbol{U}=\left\{\frac{\boldsymbol{h}-\mathbf{1}\mathrm{Tr}[\boldsymbol{h}]}{\sqrt{\left(\frac{1}{2}\mathrm{Tr}[\boldsymbol{h}]\right)^2-\det[\boldsymbol{h}]}}+\boldsymbol{\sigma}_z\right\}\times\frac{1}{\sqrt{2\left(1+\frac{(h_{11}-h_{22})/2}{\sqrt{\left(\frac{1}{2}\mathrm{Tr}[\boldsymbol{h}]\right)^2-\det[\boldsymbol{h}]}}\right)}}\tag{I.15}$$

Code

```
      program main
      implicit none
      real(8)   ,parameter :: pi=4.d0*atan(1.d0)
      complex(8),parameter :: ci=(0.d0,1.d0)
      complex(8),parameter :: cone=(1.d0,0.d0)
      real(8)   ,parameter :: ev=1.d0/27.211d0
      real(8)   ,parameter :: femtosecond=31.34137d0
      real(8)   ,parameter :: duration=50.d0*femtosecond
      real(8)   ,parameter :: level=1.d0*ev
      integer(4),parameter :: type=3
      character(64)        :: outfile='dat'
      integer(4),parameter :: nfilout=10
      integer(4),parameter :: nfilo1=12
      integer(4),parameter :: nfilo2=11
      integer(4)           :: np,ip
      real(8)              :: omega
      real(8)              :: amplitudex
      real(8)              :: amplitude
      complex(8)           :: psi(2)
      real(8)              :: scale
      character(64),parameter :: fmt1='(30("."),t1,a,t31,2f10.5)'
!     ******************************************************************
      write(*,fmt=fmt1)'duration(fs)',duration/femtosecond
      write(*,fmt=fmt1)'level(ev)',level/ev
```

```
!
      if(type.eq.1) then
        open(nfilout,file=outfile)
         omega=level+1.d-3
         amplitude=1.d-1
        np=11
        do ip=1,np
          scale=2.d0*real(ip-1,kind=8)/real(np-1,kind=8)-1.d0
          scale=1.d0+0.005d0*scale
          psi(1)=cone
          psi(2)=(0.d0,0.d0)
          call simulate(nfilout,level,omega*scale,duration,amplitude,psi)
!          print*,'scale',scale,omega*scale*abs(psi(2))**2/ev
        enddo
        stop
      else if(type.eq.2) then
         open(nfilo1,file='pfofamp_strong.dat')
         omega=level+2.d-3
         amplitude=1.d-1
         np=10000
         do ip=1,np
          scale=real(ip-1,kind=8)/real(np-1,kind=8)
          psi(1)=cone
          psi(2)=(0.d0,0.d0)
          call simulate(0,level,omega,duration,amplitude*scale,psi)
          write(nfilo1,fmt='(f10.5,4e14.4)')(scale*amplitude)**2,abs(psi(2))**2
        enddo
        close(nfilo1)
        stop
!
!     ======================================================================
!     == plot trajectories for different intensities
!     ======================================================================
      else if(type.eq.3) then
         open(nfilo2,file='pfofsqamp_weak.dat')
         amplitudex=0.3d0*eV
         omega=level+2.d-2*ev
         write(*,fmt=fmt1)'amplitudex(ev)',amplitudex/ev
         write(*,fmt=fmt1)'omega(ev)',omega/ev
         np=60
         do ip=1,np
          amplitude=amplitudex*real(ip,kind=8)/real(np,kind=8)
          write(outfile,*)nint(amplitude/ev*1.d+3)
          outfile='tra_amp'//trim(adjustl(outfile))//'milliev.dat'
          open(nfilout,file=outfile)
          psi(1)=cone
          psi(2)=(0.d0,0.d0)
          call simulate(nfilout,level,omega,duration,amplitude,psi)
          close(nfilout)
          write(nfilo2,fmt='(4e14.4)')(amplitude/ev)**2,abs(psi(2))**2
        enddo
        close(nfilo2)
        stop
      else
```

```
        stop 'type not recognized'
      end if
      stop
      end
!
!     ...1.........2.........3.........4.........5.........6.........7.........8
      subroutine simulate(nfil,deltae,omega,duration,amplitude,psi)
!     ************************************************************************
!     ** simulates the wave function of a two-state system under a          **
!     ** time-dependent perturbation in the form of a gaussian pulse        **
!     **                                                                    **
!     ** for nfil>0 the trajectory t,psi_2(t) is written to specified file. **
!     ************************************************************************
      implicit none
      integer(4),intent(in)   :: nfil
      real(8)   ,intent(in)   :: deltae    ! spacing of energy levels
      real(8)   ,intent(in)   :: omega     ! frequency of the pulse
      real(8)   ,intent(in)   :: duration  ! full width half maximum of the pulse
      real(8)   ,intent(in)   :: amplitude ! amplitude of the perturbation
      complex(8),intent(inout):: psi(2)    ! wave-function coefficients
      real(8)   ,parameter    :: pi=4.d0*atan(1.d0)
      real(8)   ,parameter :: femtosecond=31.34137d0
      logical(4)              :: tpr
      real(8)                 :: dt        ! time step
      real(8)                 :: t0        ! [-t0,t0] is the simulation time
      real(8)                 :: time      ! current time
      real(8)                 :: svar
      real(8)                 :: cw
      complex(8)              :: ham(2,2)  ! hamiltonian
!     ************************************************************************
      tpr=nfil.gt.0
      dt=(2.d0*pi/omega)*1.d-2 ! time step is small for trajectory plots
      t0=3.d0*duration
      time=-t0
      cw=duration/(2.d0*sqrt(log(2.d0)))
      if(tpr)write(nfil,fmt='(10f20.5)')time/femtosecond,0.d0,0.d0
      do
        time=time+dt
        if(time.gt.t0) exit
        ham(1,1)=(0.d0,0.d0)
        ham(2,2)=cmplx(deltae,0.d0)
        svar=amplitude*cos(omega*time)*exp(-0.5d0*(time/cw)**2)/(pi*cw)**0.25d0
        ham(1,2)=cmplx(svar)
        ham(2,1)=conjg(ham(1,2))
        call propagate(dt,ham,psi)
        if(tpr)write(nfil,fmt='(10f20.10)')time/femtosecond,abs(psi(2))**2,svar
      enddo
      if(tpr)write(nfil,fmt='(10f20.10)')time/femtosecond,0.d0,0.d0
      return
      end
!
!     ...1.........2.........3.........4.........5.........6.........7.........8
      subroutine propagate(t,h,psi)
!     ************************************************************************
```

```
!       ** propagate wave function psi under the Hamilton operator H forward    **
!       ** by the time interval t                                               **
!       **********************************************************************
        implicit none
        real(8)   ,intent(in)    :: t              ! time interval
        complex(8),intent(in)    :: h(2,2)         ! hamiltonian
        complex(8),intent(inout) :: psi(2)         ! wave function coefficients
        complex(8),parameter     :: ci=(0.d0,1.d0) ! sqrt(-1)
        logical(4),parameter     :: ttest=.false.
        real(8)                  :: trh            ! tr(h)
        real(8)                  :: deth           ! det(h)
        real(8)                  :: de
        complex(8)               :: u(2,2)
        real(8)                  :: svar
        character(64)            :: fmt1
!       **********************************************************************
        trh =real(h(1,1)+h(2,2),kind=8)
        deth=real(h(1,1)*h(2,2)-h(1,2)*h(2,1),kind=8)
        de  =sqrt((trh/2.d0)**2-deth)
        u=h
        u(1,1)=u(1,1)-0.5d0*trh
        u(2,2)=u(2,2)-0.5d0*trh
        u=u/de
        svar=sin(de*t)
        u=-ci*svar*u
        svar=cos(de*t)
        u(1,1)=-u(1,1)+svar
        u(2,2)=-u(2,2)+svar
        u=u*exp(-ci*trh/2.d0*t)
!       == propagate wave function
        psi=matmul(u,psi)
!
!       ======================================================================
!       ==  test unitarity  (test destroys u)                                ==
!       ======================================================================
        if(ttest) then
          u=matmul(u,transpose(conjg(u)))
          u(1,1)=u(1,1)-(1.d0,0.d0)
          u(2,2)=u(2,2)-(1.d0,0.d0)
          svar=sqrt(sum(abs(u)**2))
          if(svar.gt.1.d-6) then
            fmt1='(30("."),t1,a,t31,10f10.5)'
            write(*,fmt='(80("="),t10," unitarity test failed ")')
            write(*,*)'dev=',svar
            write(*,fmt=fmt1)'time',t
            write(*,fmt=fmt1)'h',h
            write(*,fmt=fmt1)'u*udagger',u
            stop
          end if
        end if
        return
        end
```

Appendix J

Propagator for time dependent Hamiltonian

We define the propagator such that it transforms the initial wave function into the time dependent wave function

$$|\Psi(t)\rangle = \hat{U}(t, t_0)|\Psi(t_0)\rangle \tag{J.1}$$

We look for a compact functional form for the propagator. We start from the Schrödinger equation, where we express the state by the propagator and the initial state.

$$\begin{aligned} i\hbar\partial_t|\psi(t)\rangle &= \hat{H}(t)|\psi(t)\rangle \\ \Rightarrow\ i\hbar\partial_t\hat{U}(t, t_0)|\psi(t_0)\rangle &= \hat{H}(t)\hat{U}(t, t_0)|\psi(t_0)\rangle \end{aligned}$$

Since this equation holds for arbitrary initial states $|\psi(t_0)\rangle$, we obtain a differential equation for the propagator

$$i\hbar\partial_t\hat{U}(t, t_0) = \hat{H}(t)\hat{U}(t, t_0)$$

with the initial condition $\hat{U}(t_0, t_0) = \hat{1}$.

Integration converts the differential equation into an integral equation

$$\hat{U}(t, t_0) == 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{H}(t')\hat{U}(t', t_0)$$

This integral equation can be solved by successive iteration, so that each approximation $\hat{U}^{(n)}(t, t_0)$ contains terms up to the n-the power of the Hamiltonian.

$$\begin{aligned} \hat{U}^{(n+1)}(t, t_0) &= \hat{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{H}(t')\hat{U}^{(n)}(t', t_0) \\ \hat{U}^{(0)}(t, t_0) &= \hat{1} \\ \hat{U}^{(1)}(t, t_0) &= \hat{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt_1\ \hat{H}(t_1) \\ \hat{U}^{(2)}(t, t_0) &= \hat{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt_1\ \hat{H}(t_1)\left[1 - \frac{i}{\hbar}\int_{t_0}^{t_1} dt_1'\ \hat{H}(t_1')\right] \\ &= \hat{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt_1\ \hat{H}(t_1) + \left(-\frac{i}{\hbar}\right)^2\int_{t_0}^{t} dt_2\ \hat{H}(t_2)\int_{t_0}^{t_2} dt_1\ \hat{H}(t_1) \\ \hat{U}^{(n)}(t, t_0) &= \sum_{j=0}^{n}\left(-\frac{i}{\hbar}\right)^j\int_{t_0}^{t} dt_j\int_{t_0}^{t_j} dt_{j-1}\ldots\int_{t_0}^{t_3} dt_2\int_{t_0}^{t_2} dt_1\ \hat{H}(t_j)\hat{H}(t_{j-1})\cdots\hat{H}(t_2)\hat{H}(t_1) \end{aligned}$$

If the series of $\hat{U}^{(n)}$ converges, that is $\hat{U}^{(n+1)} = \hat{U}^{(n)}$ in the limit $n \to \infty$, we obtain

$$\hat{U}(t,t_0) = \sum_{j=0}^{\infty} \left(-\frac{i}{\hbar}\right)^j \int_{t_0}^{t} dt_j \int_{t_0}^{t_j} dt_{j-1} \ldots \int_{t_0}^{t_2} dt_1\ \hat{H}(t_j)\hat{H}(t_{j-1})\ldots\hat{H}(t_2)\hat{H}(t_1)$$

We introduce the **time-ordering operator** $\mathcal{T}$

$$\mathcal{T}\left[\hat{H}(t_1)\hat{H}(t_2)\right] = \begin{cases} \hat{H}(t_1)\hat{H}(t_2) & \text{für} \quad t_1 > t_2 \\ \hat{H}(t_2)\hat{H}(t_1) & \text{für} \quad t_2 > t_1 \end{cases}$$

The time-ordering operator does not belong to the kind of operators we are used to: "Normal" operators act on states. This operator rearranges a product of time-dependent operators. This is why I use a calligraphic letter instead of putting a hat ontop of the symbol.

The Hamiltonians in the expression for the propagator are arranged such that the time ascends from the right to the left, i.e. $t_j \geq t_{j-1} \geq \ldots \geq t_2 \geq t_1$. Now let us allow the integration to disregard this requirement, that is each integration runs from t_0 to t. Instead of the correct order of Hamilton operator all $j!$ (j-factorial) permutation occur equally. Using the time ordering operator we bring the Hamiltonians in the right order and the factor $\frac{1}{j!}$ corrects for the multiple inclusion of the products.

$$\begin{aligned}
\hat{U}(t,t_0) &= \sum_{j=0}^{\infty} \frac{1}{j!} \left(-\frac{i}{\hbar}\right)^j \int_{t_0}^{t} dt_j \int_{t_0}^{t} dt_{j-1} \ldots \int_{t_0}^{t} dt_1\ \mathcal{T}\left[\hat{H}(t_j)\hat{H}(t_{j-1})\ldots\hat{H}(t_2)\hat{H}(t_1)\right] \\
&= \mathcal{T}\left[\sum_{j=0}^{\infty} \frac{1}{j!} \left(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\ \hat{H}(t')\right)^j\right] \\
&= \mathcal{T}\left[\mathrm{e}^{-\frac{i}{\hbar}\int_{t_0}^{t}\ \hat{H}(t')}\right]
\end{aligned}$$

Thus we obtain

PROPAGATOR

$$\hat{U}(t,t_0) = \mathcal{T}\left[\mathrm{e}^{-\frac{i}{\hbar}\int_{t_0}^{t}\ \hat{H}(t')}\right] \tag{J.2}$$

Appendix K

Hamilton-Jacobi equation and semi-classical approximation

K.1 Hamilton-Jacobi approach in more than one dimensions

Editor: The material of this section should be confirmed, e.g. tested numerically.

In this section, we map the time-dependent Schrödinger equation in the semi-classical limit onto an algorithm, which uses classical trajectories with stochastic initial conditions. A generalization where the initial conditions are chosen on a grid is possible.

K.1.1 Hamilton-Jacobi equation from the Schrödinger equation

The goal is to find an approximate solution of the time-dependent Schrödinger equation

$$\left[i\hbar\partial_t - H\left(\frac{\hbar}{i}\vec{\nabla}, \vec{r}, t\right)\right]\psi(\vec{r}, t) = 0 \tag{K.1}$$

in the quasi-classical approximation $\hbar \to 0$.

I start with the semiclassical ansatz for the wave function, namely

$$\psi(\vec{r}, t) = \mathrm{e}^{\frac{i}{\hbar}S(\vec{r},t)} . \tag{K.2}$$

The action $S(\vec{r}, t)$ is complex-valued. Below, it will be expressed in terms of a Taylor expansion in $\frac{\hbar}{i}$. The real part of the action is related to the phase of the wave function, while the imaginary part determines its absolute value.

An equation for the action is obtained by inserting the Ansatz Eq. K.2 into the Schrödinger equation

$$0 = -\mathrm{e}^{-\frac{i}{\hbar}S(\vec{r},t)} \cdot \left[i\hbar\partial_t - H\left(\frac{\hbar}{i}\vec{\nabla}, \vec{r}, t\right)\right]\mathrm{e}^{\frac{i}{\hbar}S(\vec{r},t)} = \partial_t S + H\left(\vec{\nabla}_r S + \frac{\hbar}{i}\vec{\nabla}, \vec{r}, t\right) \tag{K.3}$$

Due to the gradient in the momentum argument, the right-hand side is, in principle, still an operator . However, this gradient acts only within the Hamilton operator and does not act beyond the Hamilton operator.

For a Hamiltonian that depends (at most) quadratically on the momenta[1], this results in

$$0 = \partial_t S + H\left(\vec{\nabla}_r S, \vec{r}, t\right) + \frac{1}{2}\frac{\hbar}{i}\vec{\nabla}_r \underbrace{\vec{\nabla}_p\Big|_{\vec{p}=\vec{\nabla}_r S} H(\vec{p}, \vec{r}, t)}_{\vec{\nabla}_p H = \frac{\vec{p}-q\vec{A}}{m}} \tag{K.4}$$

[1] The generality is not yet investigated.

Editor: The quadratic dependence of the Hamilton operator on the momenta is an assumption for the following derivation!! This special form for the term depending on $\hbar$ is important, because it links to the particle conservation law, that is required below. A more general derivation would be desirable.

Let me consider the Hamilton function for a charged particle in an electromagnetic field.

$$H(\vec{p},\vec{r},t)=\frac{1}{2m}\Big(\vec{p}-q\vec{A}(\vec{r},t)\Big)^2+V(\vec{r},t) \tag{K.5}$$

For a charged particle in the electromagnetic field, I obtain the following equation of motion for the action.

$$\begin{aligned}0&=-\mathrm{e}^{-\frac{i}{\hbar}S(\vec{r},t)}\cdot\left[i\hbar\partial_t-\frac{1}{2m}\Big(\frac{\hbar}{i}\vec{\nabla}-q\vec{A}(\vec{r},t)\Big)^2-V(\vec{r},t)\right]\mathrm{e}^{\frac{i}{\hbar}S(\vec{r},t)}\\&=\partial_tS+\frac{1}{2m}\Big(\vec{\nabla}S+\frac{\hbar}{i}\vec{\nabla}-q\vec{A}\Big)^2+V(\vec{r})\\&=\partial_tS+\frac{1}{2m}\Big(\vec{\nabla}S-q\vec{A}\Big)^2+V(\vec{r})+\frac{\hbar}{i}\frac{1}{2m}\Big(\vec{\nabla}_r^2S-q\vec{\nabla}\vec{A}\Big)\end{aligned} \tag{K.6}$$

Editor: While any wave function can, in principle, be represented with the exponential ansatz Eq. K.2, such a representation may nevertheless be inconvenient and numerically problematic. Imagine a plane wave modulated with an envelope function. When the envelope function changes sign, this sign change needs to be expressed by the phase. A better description would be to decompose the wave function into serveral wave packets with node-less envelope functions.

K.1.2 Taylor expansion in $\hbar$

The differential equation Eq. K.4

$$\partial_tS+H\Big(\vec{\nabla}_rS(\vec{r},t),\vec{r}\Big)+\frac{1}{2}\frac{\hbar}{i}\vec{\nabla}_r\vec{\nabla}_p\Big|_{\vec{p}=\vec{\nabla}_rS(r,t)}H(\vec{p},\vec{r})\stackrel{\text{Eq. K.4}}{=}0 \tag{K.7}$$

converts, with the Taylor expansion

$$S(\vec{r},t)=\sum_{n=0}^{\infty}\Big(\frac{\hbar}{i}\Big)^nS_n(\vec{r},t)\;, \tag{K.8}$$

into a set of equations for the increasing orders of the action $S(\vec{r},t)$ in $\hbar$. These equations can be solved sequentially. The first two equations establish the quasi-classical limit. These are the Hamilton-Jacobi equation

$$0\stackrel{\text{Eq. K.4}}{=}\partial_tS_0(\vec{r},t)+H\Big(\vec{\nabla}_rS_0(\vec{r},t),\vec{r},t\Big) \tag{K.9}$$

for the lowest order $S_0(\vec{r},t)$ of the action, which describes the phase of the wave function. The equation for $S_1(\vec{r},t)$, the lowest order $S_1(\vec{r},t)$ of the action that provides the absolute value of the wave function

$$0\stackrel{\text{Eq. K.4}}{=}\partial_tS_1(\vec{r},t)+\left(\vec{\nabla}_p\Big|_{\vec{p}=\vec{\nabla}S_0(\vec{r},t)}H(\vec{p},\vec{r})\right)\vec{\nabla}_rS_1(\vec{r},t)+\frac{1}{2}\vec{\nabla}_r\vec{\nabla}_p\Big|_{\vec{p}=\vec{\nabla}_rS_0(\vec{r},t)}H(\vec{p},\vec{r}) \tag{K.10}$$

The equations for the higher orders are obtained by forming the $\hbar$-derivative [2] of the equation Eq. K.4 for S.

---

[2]

$$\begin{aligned}0&=\partial_t\frac{\partial S(\vec{r},t)}{\partial\hbar}+\sum_j\frac{\partial H}{\partial p_j}\Big|_{\vec{p}=\vec{\nabla}_rS}\vec{\nabla}_j\frac{\partial S}{\partial\hbar}+\frac{1}{2}\vec{\nabla}_r\,\vec{\nabla}_pH(\vec{p},\vec{r})|_{\vec{p}=\vec{\nabla}_rS}+\frac{1}{2}\frac{\hbar}{i}\Big(\vec{\nabla}_r\vec{\nabla}_p\Big)\vec{\nabla}_pH\vec{\nabla}\frac{\partial S}{\partial\hbar}\\0&=\left\{\partial_t+\sum_j\frac{\partial H}{\partial p_j}\Big|_{\vec{p}=\vec{\nabla}_rS}\vec{\nabla}_r+\frac{1}{2}\frac{\hbar}{i}\Big(\vec{\nabla}_r\vec{\nabla}_p\Big)\Big(\vec{\nabla}_pH\vec{\nabla}_r\right\}\frac{\partial S}{\partial\hbar}+\frac{1}{2}\vec{\nabla}_r\vec{\nabla}_p\Big|_{\vec{p}=\vec{\nabla}_rS(\vec{r},t)}H(\vec{p},\vec{r})\end{aligned} \tag{K.11}$$

This is a recursive set of partial differential equations of first order, which provides the contributions of increasing order in $\hbar$.

K.1.3 Hamilton's equations from the Hamilton-Jacobi equation

In lowest order of $\hbar$, we obtain the **Hamilton-Jacobi equation** Eq. K.9.

$$\partial_t S_0 + H\big(\vec{\nabla} S_0, \vec{r}\big) = 0 \tag{K.12}$$

The Hamilton-Jacobi equation is the prototype of a partial differential equation of linear order.

How does one solve the Hamilton-Jacobi equation? The goal of this section is not only to present an algorithm for solving the Hamilton-Jacobi equation, but to also motivate every ingredient of this algorithm. This may be of use if one needs to adjust the algorithm.

The general idea is to determine an integral equation for the action along some trajectory. Later, we choose the trajectories that simplify the resulting expression. This will lead us to Hamilton's equations of motion.[3]

I start with an arbitrary path[4] $\vec{R}(t)$ and work out the change of the zeroth-order action S_0 along the path

$$\frac{d}{dt} S_0(\vec{R}(t), t) = \partial_t \Big|_{\vec{r}=\vec{R}(t)} S_0(\vec{r}, t) + \dot{\vec{R}}(t)\vec{\nabla}\Big|_{\vec{r}=\vec{R}(t)} S_0(\vec{r}, t) \tag{K.13}$$

When I require that $S_0(\vec{r}, t)$ satisfies the Hamilton-Jacobi equation Eq. K.9, I can replace the partial time derivative.

$$\begin{aligned}
\frac{d}{dt} S_0(\vec{R}(t), t) &= \dot{\vec{R}}(t)\vec{\nabla} S_0(\vec{R}(t), t) \overbrace{-H\Big(\vec{\nabla} S_0(\vec{R}(t), t), \vec{R}(t)\Big)}^{\partial_t S_0 \text{ (Eq. K.9)}} \\
\Rightarrow \quad S_0(\vec{R}(t), t) &= S_0(\vec{R}(t_0), t_0) + \underbrace{\int_{t_0}^{t} dt'}_{\int d\vec{R} \frac{\dot{\vec{R}}}{|\dot{\vec{R}}|^2}} \underbrace{\Big(\dot{\vec{R}}(t')\vec{\nabla} S_0(\vec{R}(t'), t') - H\Big(\vec{\nabla} S_0(\vec{R}(t'), t'), \vec{R}(t')\Big)\Big)}_{=\dot{\vec{R}}\vec{P}-H(\vec{P},\vec{R},t)=\mathcal{L}(\dot{\vec{R}},\vec{R},t)}
\end{aligned} \tag{K.14}$$

Let me introduce the momentum[5] as a new variable for gradient of the action along the trajectory

$$\vec{P}(t) \stackrel{\text{def}}{=} \vec{\nabla} S_0(\vec{R}(t), t) \tag{K.15}$$

This transforms the integral equation for the action into a system of two equations, namely

$$S_0(\vec{R}(t), t) = S_0(\vec{R}(t_0), t_0) + \int_{t_0, [R(t)]}^{t} dt' \left(\dot{\vec{R}}(t')\vec{P}(t') - H\Big(\vec{P}(t'), \vec{R}(t'), t\Big)\right) \tag{K.16}$$

$$\text{and} \quad \vec{P}(t') = \vec{\nabla} S_0(\vec{R}(t'), t') \,. \tag{K.17}$$

This set of equations holds for arbitrary trajectories $\vec{R}(t)$ irrespective of any equations of motion for the trajectory. But, in order to perform the integral for the action, I need to know $\vec{P}(t)$ along the trajectory. This, in turn, implies that I need to know the solution $\vec{S}_0(\vec{r}, t)$ before-hand.

Note that the brace does not stop the gradients.

[3] Hamilton's equations are usually obtained from the action principle in classical mechanics. Here, they result from the quantum mechanical Schrödinger equation. Thus, we encounter one of the contact points of classical and quantum physics.

[4] I denote the path in phase space with upper-case symbols $\vec{R}(t)$ and $\vec{P}(t)$, while the phase-space coordinates $\vec{r}$ and $\vec{p}$ are denoted by lower-case symbols.

[5] At this point of the theory, the momentum is not related to the physical momentum or the momentum operator. The connection to the physical quantity will be established later.

In the search for an equation of motion for the momentum $\vec{P}(t)$ along the trajectory, let me work out the time derivative of the momentum $\vec{P}(t)$

$$\begin{aligned}\frac{d\vec{P}}{dt} &= \frac{d}{dt}\vec{\nabla}_r S_0(\vec{R}(t),t) = \underbrace{\Big(\dot{\vec{R}}(t)\vec{\nabla}_r\Big)\vec{\nabla}S_0(\vec{R}(t),t)}_{\dot{\vec{R}}\vec{\nabla}_r\otimes\vec{\nabla}_r S_0} + \underbrace{\underbrace{\vec{\nabla}_r\partial_t S_0(\vec{R}(t),t)}_{-\vec{\nabla}\big|_{\vec{R}(t)} H\left(\vec{\nabla}S_0(\vec{r},t),\vec{r},t\right)\ \text{(Eq. K.9)}}}_{-(\vec{\nabla}_p H)(\vec{\nabla}_r\otimes\vec{\nabla}_r S_0)-\vec{\nabla}_r H} \\ &\stackrel{\text{Eq. }K.15}{=} -\vec{\nabla}_r H(\vec{P},\vec{R},t) + \Big(\dot{\vec{R}} - \vec{\nabla}_p H(\vec{P},\vec{R},t)\Big)\vec{\nabla}_r\otimes\vec{\nabla}_r S_0 \end{aligned} \tag{K.18}$$

In order to obtain $\vec{\nabla}_r\otimes\vec{\nabla}_r S_0$, one not only needs the action along the trajectory but also to the side of it. This implies that the integration of the action cannot be performed for each path individually, unless the prefactor $\dot{\vec{R}} - \vec{\nabla}_P H$ vanishes.

The dependence on $\vec{\nabla}\otimes\vec{\nabla}S_0$ disappears, if the trajectory obeys

$$\dot{\vec{R}} = \vec{\nabla}_P H(\vec{P},\vec{R},t) \tag{K.19}$$

$$\dot{\vec{P}} = -\vec{\nabla}_R H(\vec{P},\vec{R},t) \tag{K.20}$$

which are **Hamilton's equations of motion**.

It is noteworthy that the only physical input for the theory is Schrödinger's equation. The classical equations of motion emerged merely as a mathematical tool to propagate the quantum-mechanical wave function. Any relation to physical observables, which are implicit in the notion of physical trajectories, still need to be established within the present theory. By proceeding along this path, this theory could be used as one possible starting point to derive classical mechanics from quantum mechanics.

K.1.4 Initial conditions

The initial conditions for the trajectories are obtained directly from the initial wave function. Given shall be a wave function $\psi(\vec{r},t_0)$, from which we need to extract the gradient of S_0 as initial momentum for the trajectories.

The wave function is converted into the initial action via

$$S_0(\vec{r},t_0) \stackrel{\text{Eq. }K.2}{=} \hbar\,\mathrm{Im}\Big[\ln\Big(\psi(\vec{r},t_0)\Big)\Big] + O(\hbar^3)\ . \tag{K.21}$$

We take the real part, because the action in lowest order of $\hbar$ is purely real-valued. Furthermore, the Hamilton-Jacobi equation is a real-valued equation. Because also S_1 (and the equation for S_1) is real-valued, only the third order in $\hbar$ contributes to the higher orders.

Thus, for any initial value $\vec{R}(t_0)$, we can initiate a trajectory. The initial momentum is then obtained from the initial wave function respectively the initial action.

$$\vec{P}(t_0) = \vec{\nabla}S_0(\vec{R}(t_0),t_0) \tag{K.22}$$

These are the initial conditions for the trajectory $(\vec{R}(t),\vec{P}(t))$.

LOWEST-ORDER ACTION S_0 FROM TRAJECTORIES

The action $S_0(\vec{R}(t), t)$ can be calculated as integral along the trajectory $(\vec{R}(t), \vec{P}(t))$ as

$$S_0(\vec{R}(t), t) \stackrel{\text{Eq. } K.16}{=} S_0(\vec{R}(t_0), t_0) + \int_{t_0,[\vec{R}(t),\vec{P}(t)]}^{t} dt' \left(\dot{\vec{R}}(t')\vec{P}(t') - H(\vec{P}(t'), \vec{R}(t'), t')\right) \qquad \text{(K.23)}$$

where the trajectory obeys Hamilton's equations of motion

$$\dot{\vec{R}} \stackrel{\text{Eq. } K.19}{=} \vec{\nabla}_p H(\vec{P}, \vec{R}, t) \qquad \text{(K.24)}$$

$$\dot{\vec{P}} \stackrel{\text{Eq. } K.20}{=} -\vec{\nabla}_r H(\vec{P}, \vec{R}, t) \qquad \text{(K.25)}$$

and the initial conditions

$$\vec{P}(t_0) \stackrel{\text{Eq. } K.22}{=} \hbar\, \vec{\nabla}\text{Im}\Big[\ln\Big(\psi(\vec{R}(t_0), t_0)\Big)\Big] + O(\hbar^3)\ . \qquad \text{(K.26)}$$

and

$$S_0(\vec{r}, t_0) \stackrel{\text{Eq. } K.2}{=} \hbar\ \text{Im}\Big[\ln\Big(\psi(\vec{r}, t_0)\Big)\Big] + O(\hbar^3)\ . \qquad \text{(K.27)}$$

The initial conditions have no well-defined limit for $\hbar \to 0$. Rather, the zeroth-order and first-order actions are determined from a wave function for a finite value of $\hbar$. The two actions, S_0 and S_1, represent the initial wave function precisely. It is the equation of motion, which is approximated by ignoring the higher orders of the action in $\hbar$, which build up during the dynamics.

K.1.5 Action principle

Rather than memorizing Hamilton's equations of motion, it is often preferable to memorize the action and its stationary condition, from which, Hamilton's equations are readily obtained.

The derivation presented below proceeds in the opposite direction than usual. Here, we start from a classical equation of motion Eq. K.19 and Eq. K.20 to calculate the action $S_0(\vec{r}, t)$. Usually one starts from the action functional $\mathcal{S}[\vec{R}(t)]$, respectively $\mathcal{S}[\vec{R}(t), \vec{P}(t)]$, of trajectories to determine the equations of motion. We will show that the action principle can be regarded as a conseqence of the equations of motion.

We consider the variation of the action due to a variation of the complete path $(\vec{R}(t), \vec{P}(t))$

including the variation of the initial conditions

$$
\begin{aligned}
\delta S_0(\vec{R}(t),t) &\stackrel{\text{Eq. K.23}}{=} \delta S_0(\vec{R}(t_0),t_0)+\int_{t_0,[\vec{P}(t),\vec{R}(t)]}^{t} dt'\,\Big(\vec{P}(t')\partial_t\delta\vec{R}(t')+\delta\vec{P}(t')\partial_t\vec{R}(t')\\
&\qquad -\delta\vec{R}(t')\vec{\nabla}_r H(\vec{P}(t'),\vec{R}(t'),t')-\delta\vec{P}(t')\vec{\nabla}_p H(\vec{P}(t'),\vec{R}(t'),t')\Big)\\
&= \delta S_0(\vec{R}(t_0),t_0)+\int_{t_0,[R(t)]}^{t} dt'\,\Big\{\frac{d}{dt}\Big(\vec{P}(t')\delta\vec{R}(t')\Big)\\
&\qquad -\delta\vec{R}(t')\Big(\partial_t\vec{P}(t')+\vec{\nabla}_r H\big(\vec{P}(t'),\vec{R}(t'),t'\big)\Big)\Big)\\
&\qquad +\delta\vec{P}(t')\Big(\partial_t R(t')-\vec{\nabla}_p H\big(\vec{P}(t'),\vec{R}(t'),t'\big)\Big\}\\
&= \delta S_0(\vec{R}(t_0),t_0)+\vec{P}(t)\delta\vec{R}(t)-\vec{P}(t_0)\delta\vec{R}(t_0)\\
&\quad +\int_{t_0,[\vec{R}(t),\vec{P}(t)]}^{t} dt'\,\Big\{\delta\vec{P}(t')\underbrace{\Big(\dot{\vec{R}}(t')-\vec{\nabla}_p H\big(\vec{P}(t'),\vec{R}(t'),t'\big)\Big)}_{=0}\\
&\qquad -\delta\vec{R}(t')\underbrace{\Big(\dot{\vec{P}}(t')-\vec{\nabla}_r H\big(\vec{P}(t'),\vec{R}(t'),t'\big)\Big)}_{=0}\Big\}\\
&= \delta S_0(\vec{R}(t_0),t_0)+\vec{P}(t)\delta\vec{R}(t)-\vec{P}(t_0)\delta\vec{R}(t_0)
\end{aligned}
\tag{K.28}
$$

If we require that the first variation of the action vanishes under variations of the trajectory $\big(\vec{P}(t),\vec{R}(t)\big)$, both Hamilton's equations of motion are satisfied.

This implies that one can use the action principle for a path in phase space as condition for the "correct" path.

The requirement that the first variation with respect to changes in the momenta vanishes can also be built into the action principle in the form

$$
\begin{aligned}
S_0(\vec{R}(t),t) &= S_0(\vec{R}(t_0),t_0)+\int_{t_0,[\vec{R}(t)]}^{t} dt'\,\underset{\vec{P}(t)}{\text{stat}}\Big(\dot{\vec{R}}(t')\vec{P}(t')-H(\vec{P}(t'),\vec{R}(t'),t)\Big)\\
&= S_0(\vec{R}(t_0),t_0)+\int_{t_0,[\vec{R}(t)]}^{t} dt'\,\mathcal{L}(\dot{\vec{R}}(t'),\vec{R}(t'),t')
\end{aligned}
\tag{K.29}
$$

This is the known expression for the action with the Lagrange function defined as

$$
\mathcal{L}(\vec{v},\vec{r},t)\stackrel{\text{def}}{=}\underset{p}{\text{stat}}\Big(\vec{v}\vec{p}-H(\vec{p},\vec{r},t)\Big)
\tag{K.30}
$$

This relates the Legendre and the Hamilton functions by a Legendre transform.

K.1.6 Particle conservation

Above, we developed an algorithm for the calculation of the zeroth order of the action $S(\vec{r},t)$ in $\hbar$. The zero'th order of the action does not provide the wave function to lowest order in $\hbar$, because it only determines the phase of the wave function, but not the amplitude. Below, I will show that the first order of S in $\hbar$, which determines the amplitude of the wave function, can be obtained from the density of classical trajectories.

Let me consider a probability distribution $W(\vec{r},t)$ of real-space points, which evolve according to Hamilton equations of motion, starting with an initial condition for the momenta $\vec{P}_j(t_0)=$

$\vec{\nabla} S_0(\vec{R}_j(t_0), t_0)$. The trajectories can also be described by a velocity field[6]

$$\vec{v}(\vec{r}, t) = \vec{\nabla}_p H(\vec{p}, \vec{r}, t)\Big|_{\vec{p}=\vec{\nabla} S_0(\vec{r}, t)} \tag{K.31}$$

The velocity field is defined for a consistent set of initial conditions for all trajectories, which is obtained from a specific initial wave function.

The particle-conservation law (considering each trajectory as a particle) in a specified velocity field is

$$\begin{aligned}
\partial_t \int_\Omega d^3r\, W(\vec{r}, t) + \oint_{\partial\Omega} d\vec{A}\, \vec{v}(\vec{r}, t) W(\vec{r}, t) &= 0 \\
\partial_t W(\vec{r}, t) + \vec{\nabla}\Big(\vec{v}(\vec{r}, t) W(\vec{r}, t)\Big) &= 0 \\
\partial_t W(\vec{r}, t) + \vec{v}(\vec{r}, t)\vec{\nabla} W(\vec{r}, t) + W(\vec{r}, t)\vec{\nabla}\vec{v}(\vec{r}, t) &= 0 \\
\frac{d}{dt} W(\vec{R}(t), t) + W(\vec{R}, t)\vec{\nabla}\vec{v}(\vec{R}, t) &= 0 \\
\frac{d}{dt} \ln\Big(W(\vec{R}(t), t)\Big) &= -\vec{\nabla}\vec{v}(\vec{R}(t), t)
\end{aligned} \tag{K.32}$$

I insert the velocity field Eq. K.31 of our trajectories.

$$\frac{d}{dt} \ln\Big(W(\vec{R}(t), t)\Big) \stackrel{\text{Eq. }K.31}{=} -\vec{\nabla}_r\Big|_{\vec{r}=\vec{R}(t)} \vec{\nabla}_p\Big|_{\vec{p}=\vec{\nabla}_r S_0(\vec{r}, t)} H(\vec{p}, \vec{r}, t) \tag{K.33}$$

This, however, is nothing but the differential equation Eq. K.11 for $S_1(\vec{r}, t)$.

$$\begin{aligned}
\frac{d}{dt} S_1(\vec{R}(t), t) &= \partial_t\Big|_{\vec{r}=\vec{R}(t)} S_1(\vec{r}, t) + \underbrace{\left(\vec{\nabla}_p\Big|_{\vec{p}=\vec{\nabla}_R S_0(\vec{R}(t), t)} H(\vec{p}, \vec{R}(t))\right)}_{\dot{\vec{R}}(t)} \vec{\nabla}_r\Big|_{\vec{r}=\vec{R}(t)} S_1(\vec{r}, t) \\
&\stackrel{\text{Eq. }K.11}{=} -\frac{1}{2}\vec{\nabla}_r\Big|_{r=R(t)} \left(\vec{\nabla}_p\Big|_{\vec{p}=\vec{\nabla}_r S_0(\vec{r}, t)} H(\vec{p}, \vec{r})\right) \stackrel{\text{Eq. }K.33}{=} \frac{1}{2}\frac{d}{dt} \ln\Big(W(\vec{R}(t), t)\Big)
\end{aligned} \tag{K.34}$$

Thus we obtain the first-order action from the density of the trajectories in real space.

$$S_1(\vec{R}(t), t) = S_1(\vec{R}(0), 0) + \frac{1}{2} \ln\left(\frac{W(\vec{R}(t), t)}{W(\vec{R}(0), 0)}\right) \tag{K.35}$$

This equation for $\vec{S}$ determines the first order term S_1 of the action. It has been obtained solely from the evolution of the density of trajectories, which have been obtained from Hamilton's equation and the initial condition of a wave function. We have not used the Schrödiger equation. Nevertheless, the equation is identical to Eq. K.10 resulting from the Schrödinger equation. This establishes, that the real-space density of trajectories provides $S_1(\vec{r}, t)$ and, thus, the absolute value of the wave function in the quasi-classical limit.

[6]The velocity field $\vec{v}(\vec{r}, t)$ can only be written down once the solution for $S_0(\vec{r}, t)$ is available. This is of no concern here, because we aim for general mathematical statement, not for a practical scheme.

K.1.7 Algorithm

SEMICLASSICAL WAVE FUNCTION FROM AN ENSEMBLE OF CLASSICAL TRAJECTORIES

This implies that, we can replace the time-dependent Schrödinger equation in the quasi-classical limit by the following stochastic algorithm.

1. Create a random set of initial positions $\vec{R}_j(t_0)$ so that their density distribution

$$W(\vec{r}, t) \stackrel{\text{def}}{=} \frac{1}{\sum_j} \sum_j \delta(\vec{r} - \vec{R}_j(t)) \tag{K.36}$$

   is equal to the squared initial wave function amplitude, i.e. $W(\vec{r}, t_0) = |\psi(\vec{r}, t_0)|^2$. The index j selects one of these trajectories

2. For each initial position determine the initial momentum from

$$\vec{P}_j(t_0) = \vec{\nabla} S_0(\vec{R}_j(t_0)) = \text{Re}\Big(\frac{\hbar}{i}\nabla \ln\Big[\Psi(\vec{R}_j(t_0), t_0)\Big]\Big) \tag{K.37}$$

3. For each set of initial coordinates $\vec{R}_j(t_0), \vec{P}_j(t_0)$, integrate Hamiltons equation of motion Eqs. K.19, and K.20 to the target time t.

4. Integrate the Lagrange function along the trajectory, Eq. K.23, to obtain the zero'th order $S_0(\vec{r}, t)$ of the action.

5. Determine the density $W(\vec{r}, t)$ of the final positions $\vec{R}_j(t)$ of the trajectories. In case of interest obtain the first order $S_1(\vec{r}, t)$ of the action from this density.

6. construct the final wave function

$$\psi(\vec{r}, t) = \sqrt{W(\vec{r}, t)}\, e^{\frac{i}{\hbar} S_0(\vec{r}, t)} \tag{K.38}$$

K.1.8 Discussion

What is the use of this algorithm?

By making contact with the wave function, in particular the initial conditions, this method allows a stringent comparison of the time dependent Schrödinger equation with classical trajectories. The algorithm provides a reasonably rigorous justification for replacing the wave function by a set of trajectories.

When the electronic and atomic motion is studied simultaneously, it is often not affordable to investigate both at a fully quantum mechanical level. Therefore, one often treats the nuclei as classical particles, while the electrons are treated quantum mechanically.

Physical problems such as diffusion, reactions etc. are easily formulated in terms of classical trajectories.

K.1.9 Example: 1-d harmonic oscillator

`Editor: This example is not completed!`

$$H(p, x) = \frac{p^2}{2m} + \frac{1}{2} c x^2 \tag{K.39}$$

The equations of motion are

$$
\begin{aligned}
&\dot{p} = -cx \quad \text{and} \quad \dot{x} = \frac{p}{m} \\
\Rightarrow \quad &\ddot{x} = \frac{1}{m}\dot{p} = -\frac{c}{m}x = -\omega^2 x
\end{aligned}
\tag{K.40}
$$

where we defined $\omega = \sqrt{c/m}$.

With the initial conditions for the trajectory $x(t=0) = x_0$ and $p(t=0) = p_0$, we obtain

$$
\begin{aligned}
x(t) &= x_0 \cos(\omega t) + \frac{p_0}{m\omega}\sin(\omega t) \\
p(t) &= m\dot{x} = -m x_0 \omega \sin(\omega t) + p_0 \cos(\omega t)
\end{aligned}
\tag{K.41}
$$

Next we need to integrate the Lagrange function

$$
\begin{aligned}
S_0(x(t),t) &= S_0(x_0,0) + \int_0^t dt' \Big(p\dot{x} - H(p,x)\Big) \\
&= S_0(x_0,0) + \int_0^t dt' \Big(\frac{1}{2}m\dot{x}^2 - \frac{1}{2}cx^2\Big) \\
&= S_0(x_0,0) + \int_0^t dt' \frac{1}{2m}\Big[\Big(-m\omega x_0 \sin(\omega t') + p_0\cos(\omega t')\Big)^2 - \overbrace{m^2\omega^2}^{mc}\Big(x_0\cos(\omega t') + \frac{p_0}{m\omega}\sin(\omega t')\Big)^2\Big] \\
&= S_0(x_0,0) + \int_0^t dt' \frac{1}{2m}\Big[\Big(-m\omega x_0 \sin(\omega t') + p_0\cos(\omega t')\Big)^2 - \Big(m\omega x_0\cos(\omega t') + p_0\sin(\omega t')\Big)^2\Big] \\
&= S_0(x_0,0) + \int_0^t dt' \frac{1}{2m}\Big[\Big(p_0^2 - m^2\omega^2 x_0^2\Big)\underbrace{\Big(\cos^2(\omega t') - \sin^2(\omega t')\Big)}_{\cos(2\omega t')} - 4m\omega x_0 p_0 \underbrace{\cos(\omega t')\sin(\omega t')}_{\frac{1}{2}\sin(2\omega t')}\Big] \\
&= S_0(x_0,0) + \frac{1}{2m}\Big[\Big(p_0^2 - m^2\omega^2 x_0^2\Big)\frac{-1}{2\omega}\sin(2\omega t) - 2m\omega x_0 p_0 \frac{1}{2\omega}\Big(\cos(2\omega t) - 1\Big)\Big] \\
&= S_0(x_0,0) - \frac{1}{4m\omega}\Big[\Big(p_0^2 - m^2\omega^2 x_0^2\Big)\sin(2\omega t) + 2m\omega x_0 p_0\Big(\cos(2\omega t) - 1\Big)\Big]
\end{aligned}
\tag{K.42}
$$

Let me now specify the initial conditions. Let the iniial wave function be a displaced Gaussian with an initial momentum p_0

$$
\psi(x, t=0) = \frac{1}{Z}\mathrm{e}^{-\lambda^2(x-\bar{x})^2}\mathrm{e}^{\frac{i}{\hbar}p_0 x}
\tag{K.43}
$$

$$
S_0(x, t=0) = p_0 x \qquad \text{and} \qquad p(t=0) = \partial_x S_0(x, t=0) = p_0
\tag{K.44}
$$

$$
S_0(x(t),t) = p_0 x_0 - \frac{1}{4m\omega}\Big[\Big(p_0^2 - m^2\omega^2 x_0^2\Big)\sin(2\omega t) + 2m\omega x_0 p_0\Big(\cos(2\omega t) - 1\Big)\Big]
\tag{K.45}
$$

with

$$
x(t) = x_0\cos(\omega t) + \frac{p_0}{m\omega}\sin(\omega t) \quad \Rightarrow \quad x_0 = \frac{m\omega x(t) - p_0\sin(\omega t)}{m\omega\cos(\omega t)}
\tag{K.46}
$$

I insert x_0 into the action and obtain

$$
\begin{aligned}
S_0(x(t),t) &= p_0 \frac{m\omega x(t) - p_0 \sin(\omega t)}{m\omega \cos(\omega t)} - \frac{1}{4m\omega}\Bigg[\left(p_0^2 - m^2\omega^2\left(\frac{m\omega x(t) - p_0 \sin(\omega t)}{m\omega\cos(\omega t)}\right)^2\right)\underbrace{\sin(2\omega t)}_{2\sin(\omega t)\cos(\omega t)} \\
&\quad + 2m\omega\frac{m\omega x(t) - p_0\sin(\omega t)}{m\omega\cos(\omega t)} p_0 \underbrace{\Big(\cos(2\omega t) - 1\Big)}_{2\cos^2(\omega t)-2}\Bigg] \\
&= \frac{-1}{4m\omega}\Bigg[-4m\omega p_0 \frac{m\omega x(t) - p_0 \sin(\omega t)}{m\omega\cos(\omega t)} + \frac{p_0^2\cos^2(\omega t) - \big(m\omega x(t) - p_0\sin(\omega t)\big)^2}{\cos^2(\omega t)}\underbrace{\sin(2\omega t)}_{2\sin(\omega t)\cos(\omega t)} \\
&\quad + 2p_0\frac{m\omega x(t) - p_0\sin(\omega t)}{\cos(\omega t)}\Big(2\cos^2(\omega t) - 2\Big)\Bigg] \\
&= \frac{-1}{2m\omega\cos(\omega t)}\Bigg[-2p_0\Big(m\omega x(t) - p_0\sin(\omega t)\Big) + \left(p_0^2\cos^2(\omega t) - \big(m\omega x(t) - p_0\sin(\omega t)\big)^2\right)\sin(\omega t) \\
&\quad + 2p_0\Big(m\omega x(t) - p_0\sin(\omega t)\Big)\Big(\cos^2(\omega t) - 1\Big)\Bigg] \\
&= \frac{-1}{2m\omega\cos(\omega t)}\Bigg[-2p_0 m\omega x(t) + 2p_0^2\sin(\omega t) \\
&\quad + \Big(p_0^2\cos^2(\omega t) - m^2\omega^2x^2(t) + 2p_0\sin(\omega t)m\omega x(t) - p_0^2\sin^2(\omega t)\Big)\sin(\omega t) \\
&\quad + 2p_0 m\omega x(t)\Big(\cos^2(\omega t) - 1\Big) - 2p_0^2\sin(\omega t)\Big)\Big(\cos^2(\omega t) - 1\Big)\Bigg] \\
&= \frac{-1}{2m\omega\cos(\omega t)}\Bigg[-\sin(\omega t)\Big(m\omega x(t)\Big)^2 + 2p_0\underbrace{\Big(-1 + \sin^2(\omega t) + \cos^2(\omega t) - 1\Big)}_{-1} m\omega x(t) \\
&\quad + p_0^2\sin(\omega t)\underbrace{\Big(2 + \cos^2(\omega t) - \sin^2(\omega t) - 2\cos^2(\omega t) + 2\Big)\Big)}_{=3}\Bigg] \\
&= \frac{\sin(\omega t)\Big(m\omega x(t)\Big)^2 + 2p_0\Big(m\omega x(t)\Big) - 3p_0^2\sin(\omega t)}{2m\omega\cos(\omega t)}
\end{aligned}
\tag{K.47}
$$

`Editor: This result looks strange. It must be checked!!!`

Interestingly, the zeroth-order action S_0 for a fixed initial momentum p_0 does not depend on the initial spatial particle distribution.

K.2 Matching of WKB solution at the classical turning point

The linearized equation

$$
\left[-\frac{\hbar^2}{2m}\partial_x^2 + \underbrace{\frac{dV}{dx}\bigg|_{x=0} x}_{V(x)-E}\right]\psi(x) = 0
$$

has the form of the so-called **Airy's equation**

$$
\partial_z^2 f(z) = z f(z) \tag{K.48}
$$

Apart from the factors Airy's equation describes the Schrödinger equation for a particle in a constant field. Thus it corresponds to the free fall.

In order to arrive at Airy's equation, we perform a variable transform[7]

$$z = \left(\sqrt[3]{-\frac{2m}{\hbar^2}\left.\frac{dV}{dx}\right|_{x=0}}\right) x$$

The solutions $f(z)$ of Airy's equation, Eq. K.48, are the Airy's functions $Ai(z)$ and $Bi(z)$.

$$Ai(z) = \frac{1}{\pi}\int_0^\infty ds\, \left(\cos(\frac{1}{3}s^3 + sz\right)$$
$$Bi(z) = \frac{1}{\pi}\int_0^\infty ds\, \left(\mathrm{e}^{-\frac{1}{3}s^3+sz} + \sin(\frac{1}{3}s^3 + sz\right)$$

Their asymptotic forms far from the origin, the classical turning point, are

$$Ai(z >> 0) = \frac{1}{2\pi z^{\frac{1}{4}}}\mathrm{e}^{-\frac{2}{3}z^{\frac{3}{2}}}$$
$$Bi(z >> 0) = \frac{1}{\pi z^{\frac{1}{4}}}\mathrm{e}^{\frac{2}{3}z^{\frac{3}{2}}}$$
$$Ai(z << 0) = \frac{1}{\pi(-z)^{\frac{1}{4}}}\sin(\frac{2}{3}(-z)^{\frac{3}{2}} + \frac{\pi}{4})$$
$$Bi(z << 0) = \frac{1}{\pi(-z)^{\frac{1}{4}}}\cos(\frac{2}{3}(-z)^{\frac{3}{2}} + \frac{\pi}{4})$$

The wave function in the neighborhood of the classical turning point can be described by the Airy's functions

$$\psi(x) = aAi(x/c) + bBi(x/c)$$

For $\hbar \to 0$, that is for $z \to \infty$ this wave function must become identical with the WKB approximation of the wave function for a linear potential.

$$\psi(x) = \frac{1}{\sqrt{|p(x)|}}\left[A\mathrm{e}^{\frac{i}{\hbar}\int_0^x dx\, p(x)} + B\mathrm{e}^{-\frac{i}{\hbar}\int_0^x dx\, p(x)}\right]$$
$$p(x) = \sqrt{2m(E - V(x))} \overset{V(x)-E=\frac{dV}{dx}x}{=} \sqrt{\left(2m\left.\frac{dV}{dx}\right|_{x=0} x\right)} = \hbar c\sqrt{-z(x)}$$

where $c = z/x$.

Thus we obtain for $z > 0$

$$\begin{aligned}
\psi(x) &= \frac{1}{\sqrt{\hbar c\sqrt{-z}}}\left[A\mathrm{e}^{\frac{i}{\hbar}\frac{1}{c}\int_0^z dz'\hbar i c z'^{\frac{1}{2}}} + B\mathrm{e}^{-\frac{i}{\hbar}\frac{1}{c}\int_0^z dz'\, i\hbar c z'^{\frac{1}{2}}}\right] \\
&\overset{(-1)^{-\frac{1}{4}}=\frac{1+i}{\sqrt{2}}}{=} \frac{(1+i)\pi}{\sqrt{2\hbar c}}\frac{1}{\pi z^{\frac{1}{4}}}\left[A\mathrm{e}^{-\frac{2}{3}z^{\frac{3}{2}}} + B\mathrm{e}^{\frac{2}{3}z^{\frac{3}{2}}}\right] \\
&\to \frac{(1+i)\pi}{\sqrt{2\hbar c}}\left[2A \cdot Ai(z) + B \cdot Bi(z)\right]
\end{aligned}$$

[7]We define the new variable z by $z = cx$ with a constant c. We insert $x = z/c$ and

$$\partial_x = \frac{d}{dx} = \frac{dz}{dx}\frac{d}{dz} = c\partial_z$$

into the Schrödinger equation at the classical turning point Eq. K.48

$$\left[-\frac{\hbar^2}{2m}c^2\partial_z^2 + \left.\frac{dV}{dx}\right|_{x=0}\frac{1}{c}z\right]\psi(x) \overset{\text{Eq. }K.48}{=} 0$$

Thus the constants drop out if

$$-\frac{\hbar^2}{2m}c^2 = +\left.\frac{dV}{dx}\right|_{x=0}\frac{1}{c} \qquad \Rightarrow \qquad c = \sqrt[3]{-\frac{2m}{\hbar^2}\left.\frac{dV}{dx}\right|_{x=0}}$$

and for $z < 0$

$$
\begin{aligned}
\psi(x) &= \frac{1}{\sqrt{\hbar c\sqrt{-z}}}\left[A\mathrm{e}^{\frac{i}{\hbar}\frac{-1}{c}\int_0^{(-z)} d(-z')\,\hbar c(-z')^{\frac{1}{2}}} + B\mathrm{e}^{-\frac{i}{\hbar}\frac{-1}{c}\int_0^{(-z)} d(-z')\,\hbar c(-z')^{\frac{1}{2}}}\right] \\
&= \frac{\pi}{\sqrt{\hbar c}}\frac{1}{\pi(-z)^{\frac{1}{4}}}\left[A\mathrm{e}^{-i\frac{2}{3}(-z)^{\frac{3}{2}}} + B\mathrm{e}^{+i\frac{2}{3}(-z)^{\frac{3}{2}}}\right] \\
&= \frac{\pi}{\sqrt{\hbar c}}\frac{1}{\pi(-z)^{\frac{1}{4}}}\left[A\mathrm{e}^{i\frac{\pi}{4}}\mathrm{e}^{-i\left(\frac{2}{3}(-z)^{\frac{3}{2}}+\frac{\pi}{4}\right)} + B\mathrm{e}^{i\frac{\pi}{4}}\mathrm{e}^{+i\left(\frac{2}{3}(-z)^{\frac{3}{2}}+\frac{\pi}{4}\right)}\right] \\
&= \frac{\pi}{\sqrt{\hbar c}}\frac{1}{\pi(-z)^{\frac{1}{4}}}\left[\left(A\mathrm{e}^{i\frac{\pi}{4}} + B\mathrm{e}^{i\frac{\pi}{4}}\right)\cos\left(\frac{2}{3}(-z)^{\frac{3}{2}} + \frac{\pi}{4}\right) + i\left(A\mathrm{e}^{i\frac{\pi}{4}} + B\mathrm{e}^{i\frac{\pi}{4}}\right)\sin\left(\frac{2}{3}(-z)^{\frac{3}{2}} + \frac{\pi}{4}\right)\right] \\
&\rightarrow \frac{\pi}{\sqrt{\hbar c}}\left[\left(A\mathrm{e}^{i\frac{\pi}{4}} + B\mathrm{e}^{i\frac{\pi}{4}}\right)\cdot Ai(z) + i\left(A\mathrm{e}^{i\frac{\pi}{4}} + B\mathrm{e}^{i\frac{\pi}{4}}\right)\cdot Bi(z)\right]
\end{aligned}
$$

The coefficients of the Airy functions must be identical left and right to the turning point. This provides us with a relation for the coefficients A, B of the solution left of the barrier with those to the right of the barrier.

Matching condition for the WKB solution at the classical turning point. The WKB wave function for a potential with positive slope at the classical turning point can be written as

$$
\begin{aligned}
\Psi(x < x_{cl}) &= \frac{1}{\sqrt{|p(x)|}}\cos\left(\frac{1}{\hbar}\left[\int_x^{x_{cl}} dx' p(x')\right] - \frac{\pi}{4}\right) \\
\Psi(x > x_{cl}) &= \frac{1}{\sqrt{|p(x)|}}\mathrm{e}^{-\frac{1}{\hbar}\int dx' p(x')}
\end{aligned}
$$

and for negative slope

$$
\begin{aligned}
\Psi(x > x_{cl}) &= \frac{1}{\sqrt{|p(x)|}}\cos\left(\frac{1}{\hbar}\left[\int_{x_{cl}}^{x} dx' p(x')\right] - \frac{\pi}{4}\right) \\
\Psi(x < x_{cl}) &= \frac{1}{\sqrt{|p(x)|}}\mathrm{e}^{\frac{1}{\hbar}\int dx' p(x')}
\end{aligned}
$$

Appendix L

Optical absorption coefficient

L.1 Macroscopic theory

Source: Solid State Physics, Part II, Optical properties of Solids, M.S. Dresselhaus, `http://web.mit.edu/course/6/6.`
retrieved Nov.1, 2016.

This section aims to clarify the terminology and to show the relation between the different physical quantities. It is fairly extended because the usual notation uses different unit systems.

We start from Maxwell's equations of motion

$$\begin{aligned}\vec{\nabla}\vec{D} &= \rho \\ \vec{\nabla}\vec{B} &= 0 \\ \vec{\nabla}\times\vec{E}+\partial_t\vec{B} &= 0 \\ \vec{\nabla}\times\vec{H}-\partial_t\vec{D} &= \vec{j}\end{aligned} \tag{L.1}$$

where $\vec{D}=\boldsymbol{\epsilon}\vec{E}$ and $\boldsymbol{\mu}\vec{H}=\hat{B}$.

We couple the currents to the electric field and exploit charge conservation[1]

$$\begin{aligned}\vec{j} &= \boldsymbol{\sigma}\vec{E} \\ \partial_t\rho+\vec{\nabla}\vec{j} &= 0 \qquad \Rightarrow \qquad \partial_t\rho=-\vec{\nabla}\sigma\vec{E}\end{aligned} \tag{L.2}$$

where $\boldsymbol{\sigma}$ is the conductivity tensor. The frequency dependent conductivity tensor is also called the **optical conductivity**.

For the sake of simplicity, I limit the discussion to isotropic materials, so that we can ignore the matrix structure of the dielectric constant.

We introduce the Fourier transformed quantities,

$$Y(\vec{r},t)=Y(\vec{k},\omega)\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)} \tag{L.3}$$

which yields for $\omega\neq 0$

$$\begin{aligned}i\vec{k}\epsilon\vec{E}=\frac{\vec{k}}{\omega}\sigma\vec{E} \qquad &\Rightarrow \qquad \Big(\sigma=i\omega\epsilon \quad \text{or} \quad \vec{k}\perp\vec{E}\Big) \\ i\vec{k}\vec{B}=0 \qquad &\Rightarrow \qquad \vec{B}\perp\vec{k} \\ i\vec{k}\times\vec{E}-i\omega\vec{B}=0 \qquad &\Rightarrow \qquad \vec{B}=\frac{1}{\omega}\vec{k}\times\vec{E} \\ i\vec{k}\times\frac{1}{\mu}\vec{B}+i\omega\epsilon\vec{E}=\boldsymbol{\sigma}\vec{E} \qquad &\Rightarrow \qquad \Big(\vec{k}\vec{E}=0 \quad \text{and} \quad \vec{k}^2-\mu\epsilon\omega^2-\mu i\omega\sigma=0\Big)\end{aligned} \tag{L.4}$$

[1] Charge conservation is implicit in Maxwell's equations. By using it explicitely we obtain an over-determined equation system.

We used the bac-cab rule, which can be decomposed into a term parallel and a term perpendicular to the electric field. The one perpendicular leads to $\vec{k}\vec{E} = 0$, while the one parallel yields the dispersion relation.

The dispersion relation can be written in the form

$$\vec{k}^2 = \mu\Big(\epsilon + i\frac{1}{\omega}\sigma\Big)\omega^2 \tag{L.5}$$

We arrive at a **complex dielectric constant**,

$$\begin{aligned}\epsilon_1 + i\epsilon_2 &= \epsilon + i\frac{1}{\omega}\sigma\\ \epsilon_{r,1} + i\epsilon_{r,2} &= \epsilon_r + i\frac{1}{\epsilon_0\omega}\sigma\end{aligned} \tag{L.6}$$

where the imaginary part is determined by the conductivity.

The complex index of refraction is defined as

$$\tilde{n}(\omega) + i\tilde{k}(\omega) = \underbrace{\frac{1}{\sqrt{\epsilon_0\mu_0}}}_{c}\underbrace{\sqrt{\mu(\epsilon_1 + i\epsilon_2)}}_{1/v_g} \tag{L.7}$$

The two quantities, the **index of refraction** $\tilde{n}$ and the **extinction coefficient** $\tilde{k}$ are called **the optical constants** of the solid.(See Dresselhaus)

The conductivity σ is also called the optical conductivity. It is the imaginary part ϵ_2 of the dielectric constant times wave number ω.

Transmitted, refracted and absorbed energy

Inside the material an incident wave has the form

$$\begin{aligned}\vec{E}(\vec{r},t) &= \vec{E}_0 e^{i(\vec{k}-\omega t)}\\ \vec{B}(\vec{r},t) &= \frac{1}{\omega}\vec{k} \times \vec{E}_0 \exp i(\vec{k} - \omega t)\end{aligned} \tag{L.8}$$

with $\vec{k} = \kappa + i\lambda$

$$\vec{k}^2 = \kappa^2 - \lambda^2 + 2i\kappa\lambda \stackrel{\text{Eq. } L.5}{=} \mu\epsilon\omega^2 + i\mu\sigma\omega \tag{L.9}$$

Thus

$$\begin{aligned}2\kappa\lambda = \mu\sigma\omega \Rightarrow \kappa &= \frac{\mu\sigma\omega}{2\lambda}\\ \kappa^2 - \lambda^2 &= \mu\epsilon\omega^2\\ \Big(\frac{\mu\sigma\omega}{2\lambda}\Big)^2 - \lambda^2 - \mu\epsilon\omega^2 &= 0\\ \lambda^4 + \mu\epsilon\omega^2\lambda^2 - \Big(\frac{\mu\sigma\omega}{2}\Big)^2 &= 0\\ \Big(\lambda^2 + \frac{\mu\epsilon\omega^2}{2}\Big)^2 &= \Big(\frac{\mu\epsilon\omega^2}{2}\Big)^2 + \Big(\frac{\mu\sigma\omega}{2}\Big)^2\\ \lambda^2 &= -\frac{\mu\epsilon\omega^2}{2} \pm \sqrt{\Big(\frac{\mu\epsilon\omega^2}{2}\Big)^2 + \Big(\frac{\mu\sigma\omega}{2}\Big)^2}\\ |\lambda| &= \sqrt{\mu\epsilon}\omega\sqrt{\frac{1}{2}\Big[-1 \pm \sqrt{1 + \Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]}\end{aligned} \tag{L.10}$$

The defining equations are invariant under the transformation $\kappa' = i\lambda$, and $\lambda' = -i\kappa$. Therefore we need to pick the solutions with real λ. The others would produce solutions for the wave length $i\kappa$. We obtain[2]

$$\begin{aligned}\lambda &= \pm\sqrt{\mu\epsilon}\omega\sqrt{\frac{1}{2}\Big[-1+\sqrt{1+\Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]}\\ \kappa &= \pm\sqrt{\mu\epsilon}\omega\sqrt{\frac{1}{2}\Big[1+\sqrt{1+\Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]}\end{aligned} \tag{L.12}$$

Because the complex index of refraction is related to the complex dielectric constant, we obtain

$$|\vec{k}| = \kappa + i\lambda = \frac{\omega}{c}\Big(\tilde{n} + i\tilde{k}\Big) \tag{L.13}$$

so that

$$\begin{aligned}\tilde{k}(\omega) &= \sqrt{\frac{1}{2}\Big[-1+\sqrt{1+\Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]}\\ \tilde{n}(\omega) &= \sqrt{\frac{1}{2}\Big[1+\sqrt{1+\Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]}\end{aligned} \tag{L.14}$$

Because the intensity is proportional to the square of the light field, the intensity falls off proportional $e^{-\frac{\mu\sigma}{\sqrt{\mu\epsilon}}z}$. Note that this term cannot be identified with the absorption coefficient, because the decay of the intensity is due to reflection and absorption.

The Intensity falls of as $e^{-2\lambda z} = e^{-2\omega\sqrt{\mu\epsilon}\tilde{k}z} = e^{-a_{abs}z}$ which establishes the link to the **absorption coefficient**. [3] (See eq.1.20 of Dresselhaus).

OPTICAL ABSORPTION COEFFICIENT AND OPTICAL CONDUCTIVITY

If the material properties, i.e. the optical conductivity, the permeability and the permittivity, are known, the optical absorption coefficient can be evaluated.

$$a_{abs}(\hbar\omega) = 2\omega\sqrt{\mu\epsilon}\tilde{k} = 2\omega\sqrt{\mu\epsilon}\sqrt{\frac{1}{2}\Big[-1+\sqrt{1+\Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]} \tag{L.15}$$

This provides us with a conversion of the optical absorption coefficient and the optical conductivity.

$$\sigma = \frac{\mu\epsilon\omega}{\mu}\sqrt{\left(1+2\left(\frac{a_{abs}(\hbar\omega)}{2\omega\sqrt{\mu\epsilon}}\right)^2\right)^2 - 1} \tag{L.16}$$

---

[2] We use $(\sqrt{1+x}-1)^{-1} = \frac{1}{x}(\sqrt{1+x}+1)$.

$$\begin{aligned}\lambda &= \pm\sqrt{\mu\epsilon}\omega\sqrt{\frac{1}{2}\Big[-1+\sqrt{1+\Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]}\\ \kappa &= \pm\frac{\mu\sigma\omega}{2}\frac{1}{\omega\sqrt{\mu\epsilon}}\frac{1}{\sqrt{\frac{1}{2}\Big[-1+\sqrt{1+\Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]}} = \pm\frac{\mu\sigma\omega}{2}\frac{1}{\omega\sqrt{\mu\epsilon}}\frac{\mu\epsilon\omega}{\mu\sigma}2\sqrt{\frac{1}{2}\Big[1+\sqrt{1+\Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]}\\ &= \pm\sqrt{\mu\epsilon}\omega\sqrt{\frac{1}{2}\Big[1+\sqrt{1+\Big(\frac{\mu\sigma}{\mu\epsilon\omega}\Big)^2}\Big]}\end{aligned} \tag{L.11}$$

[3] This is argument seems superficial, because it does not consider transmission and reflection

$\frac{a_{abs}(\hbar\omega)}{2\omega\sqrt{\mu\epsilon}}$ is equal to $1/(4\pi)$ times the ratio of absorbed energy per wavelength of the light. This can be a large number so that we can choose the approximetion

$$\sigma \approx \frac{1}{\mu}\begin{cases}\mu\epsilon\omega + \frac{1}{2\omega}\cdot a^2_{abs}(\hbar\omega) & \text{for large absorption}\\ \sqrt{\mu\epsilon}\cdot a_{abs}(\hbar\omega) & \text{for small absorption}\end{cases} \tag{L.17}$$

It can be shown that the result lies between the two bounds

$$\frac{1}{2\omega}\cdot a^2_{abs}(\hbar\omega) \le \sigma \le \frac{1}{\mu}\Big(\mu\epsilon\omega + \frac{1}{2\omega}\cdot a^2_{abs}(\hbar\omega)\Big) \tag{L.18}$$

and

$$\sigma \ge \frac{1}{\mu}\begin{cases}\frac{1}{2\omega}\cdot a^2_{abs}(\hbar\omega) & \text{accurate for large } a_{abs}\\ \\ \sqrt{\mu\epsilon}\cdot a_{abs}(\hbar\omega) & \text{accurate for small } a_{abs}\end{cases} \tag{L.19}$$

The largest relative error of the last estimate is obtained for $a_{abs} = 2\omega\sqrt{\mu\epsilon}$, where both approximations meet and provide only 70 % of the correct result.

L.2 Quantum derivation

The **optical absorption coefficient** $a(\hbar\omega)$ is the ratio of energy absorbed from a light beam by a material of unit thickness compared to the incident energy.

We consider a light pulse running through the material. The transition probabilities from the ground state into the excited states are obtained from time-dependent perturbation theory. Multiplication of this transition probabilty with the excitation energy is the absorbed energy. The energy transmitted through the material is obtained by integrating the Poynting vector of the light pulse over time.

L.2.1 Light pulse

Let me set up the light pulse using the Coulomb gauge[4]

$$\vec{A}(\vec{r},t) = C\vec{e}_A \mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}g(t) \tag{L.20}$$
$$\Phi(\vec{r},t) = 0 \tag{L.21}$$

where C is a constant with the same units as the vector potential, which will be linked below to the intensity of the incident light beam. We ignore the presence of the material, that is of the charges, that would produce a scalar potential $\Phi(\vec{r},t)$. $\vec{e}_A$ is the polarization direction of the light field. $g(t)$ is a pulse shape function.

The resulting magnetic and electric fields are

$$\begin{aligned}\vec{E}(\vec{r},t) &= -\vec{\nabla}\Phi - \partial_t\vec{A} = i\omega\vec{A} = i\omega C\vec{e}_A\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}g(t)\\ \vec{B}(\vec{r},t) &= \vec{\nabla}\times\vec{A} = i\vec{k}\times\vec{A} = i|\vec{k}|C\vec{e}_k\times\vec{e}_A\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}g(t)\end{aligned} \tag{L.22}$$

where $\vec{e}_k \stackrel{\text{def}}{=} \frac{1}{|\vec{k}|}\vec{k}$ is the direction of the wave vector.

[4] see Eq. 9.6 of ΦSX:Elektrodynamik[32]

L.2.2 Poynting vector and incident energy

The energy-flux density is the Poynting vector, Eq. 12.20 of ΦSX:Elektrodynamik[32],

$$\vec{S}(\vec{r},t)=\vec{E}(\vec{r},t)\times\vec{H}(\vec{r},t) \tag{L.23}$$

To be precise, the incident energy ΔE_{inc} that passes during the time interval Δt through the area $\vec{A}$ is $\Delta E_{inc}=\vec{S}\vec{A}\Delta t$.

Before we continue, we discuss the use of complex conjugate fields in the equation given above: We need to consider that the fields are real, so that one plane wave cannot be treated individually. When evaluating the Poynting vector for a wave

$$\begin{aligned}\vec{E}'(\vec{r},t)&=\vec{E}_0\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}+\vec{E}_0^*\mathrm{e}^{-i(\vec{k}\vec{r}-\omega t)}\\ \vec{H}'(\vec{r},t)&=\vec{H}_0\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}+\vec{H}_0^*\mathrm{e}^{-i(\vec{k}\vec{r}-\omega t)}\end{aligned} \tag{L.24}$$

one obtains

$$\begin{aligned}\vec{S}'=\vec{E}'\times\vec{H}'&=\vec{E}_0\times\vec{H}_0\mathrm{e}^{2i(\vec{k}\vec{r}-\omega t)}+\vec{E}_0\times\vec{H}_0^*+\vec{E}_0^*\times\vec{H}_0+\vec{E}_0^*\times\vec{H}_0^*\mathrm{e}^{-2i(\vec{k}\vec{r}-\omega t)}\\ &=\vec{E}_0\times\vec{H}_0^*+\vec{E}_0^*\times\vec{H}_0+2\mathrm{Re}\Big[\vec{E}_0\times\vec{H}_0\mathrm{e}^{2i(\vec{k}\vec{r}-\omega t)}\Big]\end{aligned} \tag{L.25}$$

The last term drops out in the time average, so that the intensity can be divided up into the two contributions of the indivial plane waves, if we write the Poynting vector in the form with the complex conjugate magnetic field. Thus the intensity can be expressed as a sum over plane waves, while the cross terms are interference terms that average away.[5]

This yields the Poynting vector of a plane wave as

$$\begin{aligned}\vec{S}&=\vec{E}\times\vec{H}^*=\frac{1}{\mu}\vec{E}\times\vec{B}^*\\ &\overset{\text{Eq. 13.38}}{=}\frac{1}{\mu}\Big(i\omega C\vec{e}_A\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}g(t)\Big)\times\Big(-i|\vec{k}|C^*\vec{e}_k\times\vec{e}_A\mathrm{e}^{-i(\vec{k}\vec{r}-\omega t)}g(t)\Big)\\ &=\frac{1}{\mu}\omega|\vec{k}||C|^2\underbrace{\Big(\vec{e}_A\times\Big(\vec{e}_k\times\vec{e}_A\Big)\Big)}_{\vec{e}_k-\vec{e}_A(\vec{e}_A\vec{k})}|g(t)|^2\\ &\overset{\vec{e}_k\perp\vec{e}_A}{=}\frac{1}{\mu}\omega|\vec{k}||C|^2\vec{e}_k|g(t)|^2\end{aligned} \tag{L.26}$$

We use the bac-cab identity $\vec{a}\times(\vec{b}\times\vec{c})=\vec{b}(\vec{a}\vec{c})-\vec{c}(\vec{a}\vec{b})$.

The energy flux, i.e. energy that flows through an area, is obtained by the integrating over time.

$$\frac{1}{A}E_{inc}=\int dt\vec{e}_k\vec{S}(t)=\frac{1}{\mu}\omega|\vec{k}||C|^2\int dt\,|g(t)|^2 \tag{L.27}$$

L.2.3 The interaction between light and electrons

The Hamiltonian for an electron (with charge $q_e=-e$ in an electromagnetic field is

$$\hat{H}=\frac{\Big(\hat{\vec{p}}-q_e\vec{A}(\hat{\vec{r}},t)\Big)^2}{2m_e}+q_e\Phi(\hat{\vec{r}},t) \tag{L.28}$$

[5] It needs to be shown whether this form of the Poynting vector results from Noether's theorem, which defines the Poynting vector.

Thus the perturbation due to the vector potential in the Coulomb gauge is in linear order

$$\begin{aligned}
\hat{W}(t) &= \frac{e}{2m_e}\left(\hat{\vec{p}}\hat{\vec{A}}+\hat{\vec{A}}\hat{\vec{p}}\right)+O(\vec{A}^2) \\
&= \frac{e}{2m_e}C\vec{e}_A g(t)\left(\hat{\vec{p}}\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}+\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}\hat{\vec{p}}\right)+O(\vec{A}^2) \\
&= \frac{e}{2m_e}C\vec{e}_A\mathrm{e}^{i\vec{k}\vec{r}}\left(\hbar\vec{k}+2\hat{\vec{p}}\right)\mathrm{e}^{-i\omega t}g(t)+O(\vec{A}^2) \\
&\stackrel{\vec{e}_A\perp\vec{k}}{=} \frac{e}{m_e}C\mathrm{e}^{i\vec{k}\vec{r}}\vec{e}_A\hat{\vec{p}}\mathrm{e}^{-i\omega t}g(t)+O(\vec{A}^2) \\
&= \hat{W}_0^\dagger\mathrm{e}^{i\omega t}g(t)
\end{aligned} \tag{L.29}$$

with

$$\hat{W}_0^\dagger=\frac{e}{m_e}C\vec{e}_A\mathrm{e}^{i\vec{k}\vec{r}}\hat{\vec{p}} \tag{L.30}$$

L.2.4 Transition probability and absorbed energy

The energy absorbed from a light pulse is the transition probability to an excited state times the excitation energy.

The transition probability $P_{f\leftarrow i}$ from an initial state $|i\rangle$ to a final state $|f\rangle$ can be evaluated perturbation theory, which leads to Fermi's golden rule Eq. 13.49 stated on p. 214

$$P_{f\leftarrow i}=\frac{2\pi}{\hbar}\left|\langle f|\hat{W}|i\rangle\right|^2\delta(E_f-E_i-\hbar\omega)\int dt\;|g(t)|^2 \tag{L.31}$$

where the perturbation has the form

$$\hat{W}_1(t)=\hat{W}_0\mathrm{e}^{i\omega t}g(t) \tag{L.32}$$

with pulse-shape function $g(t)$.

$|i\rangle$ and $|f\rangle$ are many-particle states. The initial state $|i\rangle$ is the ground state of the system, while the final state $|f_{u,o}\rangle=\hat{c}_u^\dagger\hat{c}_o|i\rangle$ has an additional electron-hole pair[6]. The electron-hole pair has one additional electron in the former unoccupied one-particle state $|\psi_u\rangle$, and the hole is formed by removing an electron from the formerly occupied state $|\psi_o\rangle$. The sum over final many-particle states can be replaced by a double sum over pairs of one-particle states. The matrix elements can be expressed in terms of one-particle states

$$\langle f_{u,o}|\hat{W}|i\rangle=\langle\psi_u|\hat{W}|\psi_o\rangle f(\epsilon_o)\Big(1-f(\epsilon_u)\Big) \tag{L.33}$$

In the independent-particle picture, we obtain the probability to create an electron-hole pair as

$$P_{f_{o,u}\leftarrow i}=\frac{2\pi}{\hbar}|\langle u|\hat{W}|o\rangle|^2 f(\epsilon_o)\Big(1-f(\epsilon_u)\Big)\delta(\epsilon_u-\epsilon_o-\hbar\omega)\int dt\;|g(t)|^2 \tag{L.34}$$

The absorbed energy is then obtained as

$$\begin{aligned}
E_{abs} &= \sum_f(E_f-E_i)P_{f\leftarrow i}=\hbar\omega\frac{2\pi}{\hbar}\sum_f\left|\langle f|\hat{W}_0|i\rangle\right|^2\delta(E_f-E_i-\hbar\omega)\int dt\;|g(t)|^2 \\
&\stackrel{\text{Eq. }L.30}{=} \hbar\omega\frac{2\pi}{\hbar}\sum_f\left|\left\langle f\left|\frac{e}{m_e}C\mathrm{e}^{i\vec{k}\vec{r}}\vec{e}_A\hat{\vec{p}}\right|i\right\rangle\right|^2\delta(E_f-E_i-\hbar\omega)\int dt\;|g(t)|^2 \\
&= \hbar\omega\frac{2\pi}{\hbar}\frac{e^2C^2}{m_e^2}\sum_f\left|\left\langle i\left|\mathrm{e}^{i\vec{k}\vec{r}}\vec{e}_A\hat{\vec{p}}\right|f\right\rangle\right|^2\delta(E_f-E_i-\hbar\omega)\int dt\;|g(t)|^2 \\
&= \hbar\omega\frac{2\pi}{\hbar}\frac{e^2C^2}{m_e^2}\sum_f\left|\left\langle o\left|\mathrm{e}^{i\vec{k}\vec{r}}\vec{e}_A\hat{\vec{p}}\right|u\right\rangle\right|^2 f(\epsilon_o)\Big(1-f(\epsilon_u)\Big)\delta(\epsilon_u-\epsilon_o-\hbar\omega)\int dt\;|g(t)|^2
\end{aligned}$$

[6] To keep things simple, I am limiting the discussion to one-particle excitations, which excludes excitons, plasmons and other many-particle excitations.

L.2.5 Absorption coefficient

Now we combine this result Eq. L.35 with that for the incident energy, Eq. L.27, to obtain the absorption coefficient

$$a(\hbar\omega) \stackrel{\text{def}}{=} \frac{E_{abs}/\Omega}{E_{inc}/A} \tag{L.35}$$

The absorption coefficient is the ratio of absorbed energy per volume per incident light energy passing through the absorbing material. It is an inverse length, which describes the exponential decay of the incident intensity.

Let me now evaluate the ratio of absorbed energy and incident energy.

$$\begin{aligned}
\frac{E_{abs}}{E_{inc}} &= \frac{2\pi\omega\frac{e^2C^2}{m_e^2}\sum_f \left|\left\langle o\middle|e^{i\vec{k}\vec{r}}\vec{e}_A\hat{\vec{p}}\middle|u\right\rangle\right|^2 f(\epsilon_o)\Big(1-f(\epsilon_u)\Big)\delta(\epsilon_u-\epsilon_o-\hbar\omega)\int dt\,|g(t)|^2}{A\frac{1}{\mu}\omega|\vec{k}||C|^2\int dt\,|g(t)|^2} \\
&= \frac{8\pi^2e^2}{A4\pi\epsilon_0c^2m_e^2|\vec{k}|}\sum_{o,u}\left|\left\langle o\middle|e^{i\vec{k}\vec{r}}\vec{e}_A\hat{\vec{p}}\middle|u\right\rangle\right|^2 f(\epsilon_o)\Big(1-f(\epsilon_u)\Big)\delta(\epsilon_u-\epsilon_o-\hbar\omega) \\
&= \frac{\Omega}{A}\frac{1}{4\pi\epsilon_0|\vec{k}|}\left(\frac{4\pi e}{m_ec}\right)^2\frac{1}{2}\sum_{o,u}\frac{1}{\Omega}\left|\left\langle o\middle|e^{i\vec{k}\vec{r}}\vec{e}_A\hat{\vec{p}}\middle|u\right\rangle\right|^2 f(\epsilon_o)\Big(1-f(\epsilon_u)\Big)\delta(\epsilon_u-\epsilon_o-\hbar\omega)
\end{aligned} \tag{L.36}$$

where we used $c = 1/\sqrt{\epsilon_0\mu_0}$.

The resulting absorption coefficient in the form quoted elsewhere is

$$a(\hbar\omega) = \frac{8\pi^2e^2}{4\pi\epsilon_0m_e^2c\omega}\frac{1}{\Omega}\sum_{o,u\in\Omega}\left|\langle u|e^{i\vec{k}\vec{r}}\vec{e}_A\vec{p}|o\rangle\right|^2 f(\epsilon_o)\Big(1-f(\epsilon_u)\Big)\delta(\epsilon_u-\epsilon_o-\hbar\omega) \tag{L.37}$$

This results compares, up to a factor 2 with Eq. 5.31 in Dresselhaus, when I use $4\pi\epsilon_0 = \tilde{n}$. It needs to be checked whether this is spin-degeneracy, which does not show up in my derivation, because I sum over spin-orbitals and not only spatial orbitals. It is not yet clear to me whether Dresselhaus' c is the speed of light in vacuum or in matter.

The expression given here does not make the long-wavelength limit: in the long-wavelength limit, the factor $e^{i\vec{k}\vec{r}}$ is absent. The reason for not doing the long-wave length limit is to maintain dipole-forbidden transitions, which, however, may be quadrupole allowed. Quadrupole allowed transitions are responsible for the color of gem-stones such as ruby, which absorb in the visible spectral range due to dipole-forbidden transitions between d-states on the same atom. Nevertheless, dipole-forbidden transitions are much weaker than dipole allowed ones.

In our equation, we need to consider positive and negative frequencies. The negative frequencies account for the complex conjugate of a plane wave conisidered here. The absorption coefficient for both positive and negative frequencies is the average of the two absorption coefficients.

L.2.6 Dipole operator

The expression for the absorption coefficient given above is convenient for solids. It has the disadvantage that the matrix elements are not easy to comprehend. For finite systems such as atoms and molecules, the dipole operator is very convenient. Here we transform the expresssion for the absorption coeffcient into the form of a dipole operator. **Here we leave out the factor $e^{i\vec{k}\vec{r}}$, which is responsible for the dipole-forbidden transition probability.**

We use a "trick":

$$\vec{e}_A = \frac{i}{\hbar}[\vec{p}, \vec{e}_A\vec{r}]_- \tag{L.38}$$

which yields

$$
\begin{aligned}
\vec{e}_A\vec{p} &= \frac{1}{2}\left(\vec{e}_A\vec{p}+\vec{p}\vec{e}_A\right) = \frac{i}{2\hbar}\left(\vec{p}[\vec{p},\vec{e}_A\vec{r}]_- + [\vec{p},\vec{e}_A\vec{r}]_-\vec{p}\right) = \frac{i}{2\hbar}\left(\vec{p}^2\vec{e}_A\vec{r} - \vec{e}_A\vec{r}]_-\vec{p}^2\right) \\
&= \frac{im_e}{\hbar}\left[\hat{H}_0,\vec{e}_A\vec{r}\right]_-
\end{aligned}
\tag{L.39}
$$

where $\hat{H}_0 = \frac{\hat{\vec{p}}^2}{2m_e} + V(\vec{r})$ is the unperturbed Hamilton operator.

The matrix elements are

$$
\langle u|\vec{e}_A\vec{p}|o\rangle = \frac{im_e}{\hbar}\Big\langle u\Big|\left[\hat{H}_0,\vec{e}_A\vec{r}\right]_-\Big|o\Big\rangle = \frac{im_e}{\hbar}(\epsilon_u-\epsilon_o)\big\langle u\big|\vec{e}_A\vec{r}\big|o\big\rangle
\tag{L.40}
$$

Thus, the momentum matrix elements in the absorption coeffcient Eq. L.37 have been replaced by those of the dipole operator. Secondly, the factor $\hbar\omega = \epsilon_o - \epsilon_u$ cancels a similar factor in the denomnator of the absorption coefficient.

Now, we insert this result into the expression for the absorption coefficient.

$$
\begin{aligned}
a(\hbar\omega) &= \frac{8\pi^2e^2}{4\pi\epsilon_0 m_e^2 c\omega}\sum_{o,u\in\Omega}\frac{1}{\Omega}\Big|\langle u|\vec{e}_A\vec{p}|o\rangle\Big|^2 f(\epsilon_o)\Big(1-f(\epsilon_u)\Big)\delta(\epsilon_u-\epsilon_o-\hbar\omega) \\
&\stackrel{\text{Eq. L.40}}{=} \underbrace{\left(\frac{e^2}{4\pi\epsilon_0\hbar c}\right)}_{\alpha}\frac{8\pi^2}{\hbar\omega}\sum_{o,u\in\Omega}(\epsilon_u-\epsilon_o)^2\frac{1}{\Omega}\Big|\big\langle u\big|\vec{e}_A\vec{r}\big|o\big\rangle\Big|^2 f(\epsilon_o)\Big(1-f(\epsilon_u)\Big)\delta(\epsilon_u-\epsilon_o-\hbar\omega) \\
&= 8\pi^2\alpha\hbar\omega\sum_{o,u\in\Omega}\frac{1}{\Omega}\Big|\langle u|\vec{e}_A\vec{r}|o\rangle\Big|^2 f(\epsilon_o)\Big(1-f(\epsilon_u)\Big)\delta(\epsilon_u-\epsilon_o-\hbar\omega)
\end{aligned}
\tag{L.41}
$$

where $\alpha = e^2/(4\pi\epsilon_0\hbar c)$ is the **fine-structure constant**. The unit of the absorption coefficient is 1/length.[7]

L.2.7 Absorption coefficient and joint density of states

The sum over states can be performed using the density of states

$$
D_{\alpha,\beta}(\epsilon) = \sum_n \langle\pi_\alpha|\psi_n\rangle\delta(\epsilon-\epsilon_n)\langle\psi_n|\pi_\beta\rangle
\tag{L.42}
$$

where we use the projector functions $\langle\pi_\alpha|$ for an basis orbital $|\chi_\alpha\rangle$, which are defined by

$$
|\psi_n\rangle = \sum_\alpha |\chi_\alpha\rangle\langle\pi_\alpha|\psi_n\rangle
\tag{L.43}
$$

$$
\begin{aligned}
a(\hbar\omega) &= 8\pi^2\alpha\,\frac{1}{\Omega}\sum_{\alpha,\beta,\gamma,\delta\in\Omega}\langle\chi_\beta|\vec{e}_A\vec{r}|\chi_\gamma\rangle\langle\chi_\delta|\vec{e}_A\vec{r}|\chi_\alpha\rangle \\
&\times\hbar\omega\int d\epsilon\; D_{\alpha,\beta}(\epsilon)D_{\gamma,\delta}(\epsilon+\hbar\omega)f(\epsilon)\Big(1-f(\epsilon+\hbar\omega)\Big)
\end{aligned}
\tag{L.44}
$$

L.3 Peierls substitution

Peierls substitution[51] (The term has been coined by Hofstadter[52]) is an alternative way to study optical excitations, which does not directly work with the dipole operator. This is advantageous for

[7] The delta function has unit 1/energy.

solids, where the dipole operator is ill-defined due to the boundary conditions. Peierls substitution constructs a basisset, which depends on the vector potential. The Hamiltonian in the new basisset is similar to the original Hamiltonian, but it includes a phase factor, the **Peierls phase**, into the hopping matrix elements between different sites.

The Peierls substitution method has been developed for time-independent magnetic fields. Here, it is generalized to time dependent problems.

Coulomb gauge: Consider the Hamiltonian for electrons in an electromagnetic field described by a time-dependent vector potential $\vec{A}(\vec{r},t)$ in the Coulomb gauge, i.e. $\vec{\nabla}\vec{A}=0$. (see also ΦSX: Elektrodynamics[32], section 9.3.1). In the Coulomb gauge, the scalar potential is due to the instantaneous charge distribution, and the vector potential obeys

$$\begin{aligned}\Phi(\vec{r},t) &= \int d^3r'\ \frac{\rho(\vec{r'},t)}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\\ \left(\vec{\nabla}^2-\frac{1}{c^2}\partial_t^2\right)\vec{A} &= -\mu_0\vec{j}+\frac{1}{c^2}\vec{\nabla}\Phi\\ \vec{\nabla}\vec{A} &= 0\end{aligned} \tag{L.45}$$

We consider here an electrostatic potential $\Phi(\vec{r}) =: V(\vec{r})$ from a static charge distribution of a material, and an electromagnetic wave described by a vector potential.

$$\vec{A}(\vec{r},t)=\sum_{\vec{k}}\left(\vec{A}_0(\vec{k})e^{i(\vec{k}\vec{r}-c|\vec{k}|t)}+\vec{A}_0^*(\vec{k})e^{-i(\vec{k}\vec{r}-c|\vec{k}|t)}\right)\qquad \text{with } \vec{A}_0(\vec{k})\perp\vec{k} \tag{L.46}$$

With this ansatz, we exclude the emission of light due to induced currents.

Basisset: The Hamiltonian for the electrons with the vector potential included has the form

$$\hat{H}(t):=\frac{1}{2m_e}\left(\vec{p}-q\vec{A}(\vec{r},t)\right)^2+V(\vec{r})\ , \tag{L.47}$$

With $V(\vec{r})$, I denote the static potential acting on the electrons.

I denote the Hamiltonian in the absence of a time-dependent electromagnetic field with

$$\hat{H}_0:=\frac{\vec{p}^2}{2m_e}+V(\vec{r}) \tag{L.48}$$

Starting from some atom-centered local basisset $\{|\chi_\alpha\rangle\}$, we choose a new basisset $\{|\tilde{\chi}_\alpha\rangle\}$ with orbitals that depend explicitly on the vector potential, namely

$$\langle\vec{r}|\tilde{\chi}_\alpha(t)\rangle=\exp\left[\frac{i}{\hbar}q\int_{\vec{R}_\alpha}^{\vec{r}}d\vec{r'}\ \vec{A}(\vec{r'},t)\right]\langle\vec{r}|\chi_\alpha\rangle \tag{L.49}$$

The orbital $|\chi_\alpha\rangle$ is centered at the atom position $\vec{R}_\alpha$. The line integral of the vector potential in Eq. L.49 is path dependent. To make the definition unique, we choose a straight line from $\vec{R}_\alpha$ to $\vec{r}$ as integration line. Note, that the orbitals $|\tilde{\chi}_\alpha\rangle$ have an explicit time dependence built in.

Hamiltonian

Next, we wish to solve the time-dependent Schrödinger equation for the wave function

$$|\psi(t)\rangle=\sum_\beta|\tilde{\chi}_\beta(t)\rangle c_\beta(t) \tag{L.50}$$

We need to consider the time and space dependence of the orbitals $|\tilde{\chi}_\beta\rangle$ in order to obtain the equation of motion for the coefficients

$$\Big(i\hbar\partial_t - \hat{H}(t)\Big)\underbrace{\sum_\beta |\tilde{\chi}_\beta(t)\rangle c_\beta(t)}_{|\psi(t)\rangle} = 0$$

$$\Rightarrow \qquad \sum_\beta\bigg(\langle\tilde{\chi}_\alpha|\tilde{\chi}_\beta\rangle i\hbar\partial_t - \langle\tilde{\chi}_\alpha|\hat{H}(t)|\tilde{\chi}_\beta\rangle + \Big\langle\tilde{\chi}_\alpha\Big|i\hbar\partial_t\tilde{\chi}_\beta\Big\rangle\bigg)c_\beta(t) = 0 \qquad \text{(L.51)}$$

Let us first study, how the derivatives act on the phase factor.

-

$$\begin{aligned}\langle\vec{r}|\left(\hat{\vec{p}}-q\hat{\vec{A}}\right)|\tilde{\chi}_\alpha\rangle &\stackrel{\text{Eq. L.49}}{=} \left(\frac{\hbar}{i}\vec{\nabla}-q\vec{A}(\vec{r},t)\right)\underbrace{\exp\left[\frac{i}{\hbar}q\int_{R_\alpha}^{\vec{r}} d\vec{r'}\;\vec{A}(\vec{r'},t)\right]\langle\vec{r}|\chi_\alpha\rangle}_{\langle\vec{r}|\tilde{\chi}_\alpha(t)\rangle}\\
&= \exp\left[\frac{i}{\hbar}q\int_{R_\alpha}^{\vec{r}} d\vec{r'}\;\vec{A}(\vec{r'},t)\right]\left(q\vec{A}(\vec{r},t)+\frac{\hbar}{i}\vec{\nabla}-q\vec{A}(\vec{r},t)\right)\langle\vec{r}|\chi_\alpha\rangle\\
&= \exp\left[\frac{i}{\hbar}q\int_{R_\alpha}^{\vec{r}} d\vec{r'}\;\vec{A}(\vec{r'},t)\right]\langle\vec{r}|\vec{p}|\chi_\alpha\rangle \qquad \text{(L.52)}\end{aligned}$$

This yields

$$\begin{aligned}\hat{H}(t)|\tilde{\chi}_\alpha\rangle &\stackrel{\text{Eq. L.47}}{=} \exp\left[\frac{i}{\hbar}q\int_{R_\alpha}^{\vec{r}} d\vec{r'}\;\vec{A}(\vec{r'},t)\right]\left(\frac{\hat{\vec{p}}^2}{2m_e}+V(\vec{r})\right)|\chi_\alpha\rangle\\
&= \exp\left[\frac{i}{\hbar}q\int_{R_\alpha}^{\vec{r}} d\vec{r'}\;\vec{A}(\vec{r'},t)\right]\hat{H}_0|\chi_\alpha\rangle \qquad \text{(L.53)}\end{aligned}$$

- We take into account that the basis functions $|\tilde{\chi}_\alpha\rangle$ inherit a time dependence from the vector potential

$$\begin{aligned}-i\hbar\partial_t\big|\tilde{\chi}_\beta(t)\big\rangle &\stackrel{\text{Eq. L.49}}{=} q\int_{\vec{R}_\beta}^{\vec{r}} d\vec{r'}\;\dot{\vec{A}}(\vec{r'},t)|\tilde{\chi}_\beta(t)\rangle\\
&= \exp\left[\frac{i}{\hbar}q\int_{R_\alpha}^{\vec{r}} d\vec{r'}\;\vec{A}(\vec{r'},t)\right]\underbrace{\left[-q\int_{\vec{R}_\beta}^{\vec{r}} d\vec{r'}\;\left(\vec{E}(\vec{r'},t)-\vec{E}(\vec{r'},0)\right)\right]}_{\approx -q(\vec{r}-\vec{R}_\beta)\cdot\left(\vec{E}(t)-\vec{E}(0)\right)}|\chi_\beta(t)\rangle \qquad \text{(L.54)}\end{aligned}$$

The electric field is[8] $\vec{E}(\vec{r},t) = -\vec{\nabla}\Phi(\vec{r},t) - \partial_t\vec{A}(\vec{r},t)$. To exclude the static electric fields due to the charge distribution, we consider only dynamic electric fields, i.e. $\vec{E}(t) - \vec{E}(0)$.

When we evaluate the matrix elements for the Hamiltonian, we encounter the product of two

[8] see Eq. 9.2 in ΦSX: Electrodynamics[32].

phase factors:

$$
\begin{aligned}
&\exp\left[\frac{i}{\hbar}q\int_{\vec{R}_\alpha}^{\vec{r}} d\vec{r'}\,\vec{A}(\vec{r'},t)\right]^* \exp\left[\frac{i}{\hbar}q\int_{\vec{R}_\beta}^{\vec{r}} d\vec{r'}\,\vec{A}(\vec{r'},t)\right]\\
&=\exp\left[-\frac{i}{\hbar}q\int_{\vec{R}_\alpha}^{\vec{R}_\beta} d\vec{r'}\,\vec{A}(\vec{r'},t)\right]\exp\left[\frac{i}{\hbar}q\underbrace{\left(\int_{\vec{R}_\alpha}^{\vec{R}_\beta} d\vec{r'}\,\vec{A}(\vec{r'},t)+\int_{\vec{r}}^{\vec{R}_\alpha} d\vec{r'}\,\vec{A}(\vec{r'},t)+\int_{\vec{R}_\beta}^{\vec{r}} d\vec{r'}\,\vec{A}(\vec{r'},t)\right)}_{\mathcal{F}_{\alpha,\beta}(\vec{r})}\right]\\
&=\underbrace{\exp\left[-\frac{i}{\hbar}q\int_{\vec{R}_\alpha}^{\vec{R}_\beta} d\vec{r'}\,\vec{A}(\vec{r'},t)\right]}_{\text{Peierls phase}}\underbrace{\mathrm{e}^{\frac{i}{\hbar}q\mathcal{F}_{\alpha,\beta}(\vec{r})}}_{\approx 1}
\end{aligned}
\tag{L.55}
$$

- The factor

$$
\exp\left[-\frac{i}{\hbar}q\int_{\vec{R}_\alpha}^{\vec{R}_\beta} d\vec{r'}\,\vec{A}(\vec{r'},t)\right]
\tag{L.56}
$$

is the **Peierls phase**.

- The quantity

$$
\begin{aligned}
\mathcal{F}_{\alpha,\beta}(\vec{r}) &\stackrel{\text{def}}{=} \int_{\vec{R}_\alpha}^{\vec{R}_\beta} d\vec{r'}\,\vec{A}(\vec{r'},t)+\int_{\vec{r}}^{\vec{R}_\alpha} d\vec{r'}\,\vec{A}(\vec{r'},t)+\int_{\vec{R}_\beta}^{\vec{r}} d\vec{r'}\,\vec{A}(\vec{r'},t)\\
&=\oint_{\partial S} d\vec{r'}\,\vec{A}(\vec{r'},t)\stackrel{\text{Stokes}}{=}\int_S d\vec{a'}\,\vec{\nabla}'\times\vec{A}(\vec{r'},t)\stackrel{\text{Stokes}}{=}\int_S d\vec{a'}\,\vec{B}(\vec{r'},t)
\end{aligned}
\tag{L.57}
$$

is the magnetic flux through the triangle S, which has the corners R_α, $\vec{R}_\beta$, and $\vec{r}$. With ∂S we denote the boundary of the triangle and $d\vec{a}$ is a surface element on it. $\vec{B}=\vec{\nabla}\times\vec{A}$ is the magnetic field

For a constant magnetic field, we can evaluate $\mathcal{F}_{\alpha,\beta}(\vec{r})$ as follows: The area of the triangle is the size of the vector $\frac{1}{2}(\vec{r}-\vec{R}_\beta)\times(\vec{R}_\alpha-\vec{R}_\beta)$. This vector is perpendicular to the triangle. Thus

$$
\mathcal{F}_{\alpha,\beta}(\vec{r})=\frac{1}{2}\vec{B}\cdot\Big[(\vec{r}-\vec{R}_\beta)\times(\vec{R}_\alpha-\vec{R}_\beta)\Big]\stackrel{\text{cycl.perm.}}{=}-(\vec{r}-\vec{R}_\beta)\underbrace{\left[\frac{1}{2}\vec{B}\times(\vec{R}_\alpha-\vec{R}_\beta)\right]}_{\text{like an electric field}}
\tag{L.58}
$$

is a linear potential which vanishes in the plane spanned by the "bond axis" and the magnetic field. It acts like an effective electric field. This field is, however, neglegibly small, because it is of size "magnetic field times interatomic distance".

PEIERLS SUBSTITUTION

The time dependent coefficients of the wave function

$$\psi(\vec{r},t)=\sum_\beta \tilde{\chi}_\beta(\vec{r},t)c_\beta(t) \stackrel{\text{Eq. } L.49}{=} \sum_\beta \underbrace{\exp\left[\frac{i}{\hbar}q\int_{\vec{R}_\beta}^{\vec{r}} d\vec{r'}\;\vec{A}(\vec{r'},t)\right]\chi_\beta(\vec{r})}_{\tilde{\chi}_\beta(\vec{r},t)}\,c_\beta(t) \qquad \text{(L.59)}$$

obey a differential equation

$$\sum_\beta \tilde{O}_{\alpha,\beta}(t)i\hbar\partial_t c_\beta(t) \stackrel{\text{Eq. } L.51}{=} \sum_\beta \tilde{H}_{\alpha,\beta}(t)c_\beta(t) \qquad \text{(L.60)}$$

where the matrix elements are

$$\begin{aligned}
\tilde{H}_{\alpha,\beta}(t) &\stackrel{\text{Eq. } L.55}{=} \left\langle\tilde{\chi}_\alpha\middle|\hat{H}(t)-i\hbar\partial_t\middle|\tilde{\chi}_\beta\right\rangle\\
&= \left\langle\chi_\alpha\middle|e^{\frac{i}{\hbar}\mathcal{F}_{\alpha,\beta}(\vec{r},t)}\left(\hat{H}_0-q\int_{\vec{R}_\beta}^{\vec{r}} d\vec{r'}\;\left(\vec{E}(\vec{r'},t)-\vec{E}(\vec{r'},0)\right)\right)\middle|\chi_\beta\right\rangle\\
&\quad\times\exp\left[-\frac{i}{\hbar}q\int_{\vec{R}_\alpha}^{\vec{R}_\beta} d\vec{r'}\;\vec{A}(\vec{r'},t)\right]\\
\tilde{O}_{\alpha,\beta}(t) &= \left\langle\tilde{\chi}_\alpha\middle|\tilde{\chi}_\beta\right\rangle\\
&= \left\langle\chi_\alpha\middle|e^{\frac{i}{\hbar}\mathcal{F}_{\alpha,\beta}(\vec{r},t)}\middle|\chi_\beta\right\rangle\cdot\exp\left[-\frac{i}{\hbar}q\int_{\vec{R}_\alpha}^{\vec{R}_\beta} d\vec{r'}\;\vec{A}(\vec{r'},t)\right] \qquad \text{(L.61)}
\end{aligned}$$

where $\mathcal{F}_{\alpha,\beta}(\vec{r},t)$ is defined in Eq. L.57.

The Peierls substitution is without approximations as long as $\mathcal{F}_{\alpha,\beta}(\vec{r})$ is not set to zero. It is valid even in strong fields and and for short wave lengths. The basisset needs to be local[9], but can be non-orthogonal.

L.3.1 Approximations of Peierls substitution

Peierls substitution separates complex, but small terms out. To make profit from it, we need to make approximations.

1. Firstly, we neglect the small terms $\mathcal{F}_{\alpha,\beta}(\vec{r})$. In order to provide a intuition, let me extract the leading error term.

$$\exp\left(\frac{iq}{\hbar}\mathcal{F}_{\alpha,\beta}(\vec{r})\right)\approx 1+\frac{iq}{2\hbar}(\vec{r}-\vec{R}_\beta)\Big[(\vec{R}_\alpha-\vec{R}_\beta)\cdot\left(\vec{\nabla}\big|_{\vec{R}_\beta}\times\vec{A}\right)\Big] \qquad \text{(L.62)}$$

2. Let me restrict the onsite terms to dipole and quadupole terms

$$\int_{R_\beta}^{\vec{r}} d\vec{r'}\Big(\vec{E}(\vec{r'},t)-\vec{E}(\vec{r'},0)\Big)\approx(\vec{r}-\vec{R}_\beta)\vec{E}(\vec{r}_\beta)+\frac{1}{2}(\vec{r}-\vec{R}_\beta)\Big(\vec{\nabla}\otimes\vec{E}(\vec{r}_\beta)\Big)(\vec{r}-\vec{R}_\beta) \qquad \text{(L.63)}$$

   The first term will be responsible for dipole-allowed, on-site transitions. whereas the second will account for quadrupole allowed transitions.All further terms will be ignored.

3. Secondly, we include only onsite matrix elements of the terms above.

[9]The basisset must be localized to keep $\mathcal{F}$ small. An extended basisset such as plane waves, would not be suitable.

The resulting equations of motion are

$$
\begin{aligned}
&\sum_\beta \exp\left[-\frac{i}{\hbar}q\int_{\vec{R}_\alpha}^{\vec{R}_\beta} d\vec{r}'\;\vec{A}(\vec{r}',t)\right]\langle\chi_\alpha|\chi_\beta\rangle i\hbar\partial_t c_\beta(t)\\
&\stackrel{\text{Eq. } L.51}{=} \sum_\beta \underbrace{\left\{\exp\left[-\frac{i}{\hbar}q\int_{\vec{R}_\alpha}^{\vec{R}_\beta} d\vec{r}'\;\vec{A}(\vec{r}',t)\right]\langle\chi_\alpha|\hat{H}_0|\chi_\beta\rangle\right\}}_{\text{charge transfer excitations}} c_\beta(t)\\
&\quad+\quad \sum_\beta \delta_{R_\alpha,R_\beta}\underbrace{\left\{\;\underbrace{-q\dot{\vec{A}}(\vec{R}_\beta,t)}_{q[\vec{E}(\vec{R}_\beta,t)-\vec{E}(\vec{R}_\beta,0)]}\;\;\big\langle\chi_\alpha\big|(\vec{r}-\vec{R}_\beta)\big|\chi_\beta\big\rangle\right\}}_{\text{onsite dipole-allowed excitations}} c_\beta(t)\\
&\quad+\quad \sum_\beta \delta_{R_\alpha,R_\beta}\underbrace{\left\{-q\mathrm{Tr}\Big[\Big(\vec{\nabla}\otimes\dot{\vec{A}}(\vec{R}_\beta,t)\Big)\frac{1}{2}\big\langle\chi_\alpha\big|(\vec{r}-\vec{R}_\beta)\otimes(\vec{r}-\vec{R}_\beta)\big|\chi_\beta\big\rangle\Big]\right\}}_{\text{onsite quadrupole-allowed excitations}} c_\beta(t)
\end{aligned}
\tag{L.64}
$$

The first term describes charge transfer excitations, which involve orbitals from different sites. The second term describes onsite dipole-allowed excitations.

L.3.2 Perturbation due to the electromagnetic field: Peierls substitution

Here, we use the equations above to work out the time dependent perturbation for a light pulse.

We proceed analogous to appendix B of `https://arxiv.org/pdf/1610.07548.pdf`, to express the perturbation due to the electromagnetic field using the Peierls-substitution ansatz.

Describe the excitation using the electromagnetic field

$$
\begin{aligned}
\vec{A}(\vec{r},t) &= \vec{e}_A\frac{1}{2}\Big(A_0\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}+A_0^*\mathrm{e}^{-i(\vec{k}\vec{r}-\omega t)}\Big)g(t)\\
\Phi(\vec{r},t) &= 0
\end{aligned}
\tag{L.65}
$$

$g(t)$ is a smooth envelope function which is normalized to one. We take it as a real number. A_0 is a complex amplitude.

Now we take the long-wave-length approximation

$$
\begin{aligned}
\vec{A}(\vec{r},t) &= \vec{e}_A\frac{1}{2}\Big(A_0\mathrm{e}^{-i\omega t}+A_0^*\mathrm{e}^{i\omega t}\Big)g(t)\\
&\quad+\vec{e}_A\frac{1}{2}\Big(A_0 i\vec{k}\vec{r}\mathrm{e}^{-i\omega t}-A_0^* i\vec{k}\vec{r}\mathrm{e}^{i\omega t}\Big)g(t)+O(|\vec{k}|^2)
\end{aligned}
\tag{L.66}
$$

The Peierls phase is in the long-wave length limit

$$
\exp\left[\frac{i}{\hbar}q\int_{\vec{R}_\beta}^{\vec{R}_\alpha} d\vec{r}'\;\vec{A}(\vec{r}',t)\right]\approx\exp\left[\frac{i}{\hbar}q\vec{e}_A\left(\vec{R}_\alpha-\vec{R}_\beta\right)\mathrm{Re}\left[A_0\mathrm{e}^{i\omega t}\right]g(t)\right]
\tag{L.67}
$$

$$
\begin{aligned}
\hat{H}(t) &= \sum_{\alpha,\beta}\exp\left[\frac{i}{\hbar}q\vec{e}_A\left(\vec{R}_\alpha-\vec{R}_\beta\right)\mathrm{Re}\left[A_0\mathrm{e}^{i\omega t}\right]g(t)\right]h_{\alpha,\beta}\hat{c}^\dagger_\alpha\hat{c}_\beta\\
&= \underbrace{\sum_{\alpha,\beta}h_{\alpha,\beta}\hat{c}^\dagger_\alpha\hat{c}_\beta}_{\hat{H}_0}+\underbrace{\sum_{\alpha,\beta}\Big\{\exp\left[\frac{i}{\hbar}q\vec{e}_A\left(\vec{R}_\alpha-\vec{R}_\beta\right)\mathrm{Re}\left[A_0\mathrm{e}^{i\omega t}\right]g(t)\right]-1\Big\}h_{\alpha,\beta}\hat{c}^\dagger_\alpha\hat{c}_\beta}_{\hat{W}(t)}
\end{aligned}
\tag{L.68}
$$

The time dependent perturbation has the form.

$$\hat{W}(t)=\sum_{\alpha,\beta}\Big\{\exp\Big[\frac{i}{\hbar}q\vec{e}_A\Big(\vec{R}_\alpha-\vec{R}_\beta\Big)\,\mathrm{Re}\left[A_0\mathrm{e}^{i\omega t}\right]g(t)\Big]-1\Big\}h_{\alpha,\beta}\hat{c}^\dagger_\alpha\hat{c}_\beta \tag{L.69}$$

Only intersite hopping terms contribute to the time-dependent perturbation. For nearest neighbor hopping, only nearest neighbor terms contributes to the time-dependent perturbation.

L.3.3 Perturbation due to the electromagnetic field: Dipole operator

Editor: This section is probably redundant. Compare with section L.2.6

The Hamiltonian of an electron in an electromagnetic field is

$$\begin{aligned}\hat{H}&=\frac{(\vec{p}-q\vec{A}(\vec{r},t))^2}{2m_e}+\Phi(\vec{r},t)+V(\vec{r})\\&=\underbrace{\frac{\vec{p}^2}{2m_e}+V(\vec{r})}_{\hat{H}_0}+\underbrace{\Big[-\frac{q}{2m}\Big(\vec{p}\vec{A}+\vec{A}\vec{p}\Big)+\frac{q^2}{2m_e}\vec{A}^2+\Phi(\vec{r},t)\Big]}_{\hat{W}(t)}\end{aligned} \tag{L.70}$$

$$\begin{aligned}\frac{i}{2\hbar}[\hat{\vec{p}}^2,\hat{x}_\alpha]_-&=\frac{i}{2\hbar}\Big(\hat{\vec{p}}^2\hat{x}_\alpha\underbrace{-\hat{\vec{p}}\hat{x}_\alpha\hat{\vec{p}}+\hat{\vec{p}}\hat{x}_\alpha\hat{\vec{p}}}_{=0}-\hat{x}_\alpha\hat{\vec{p}}^2\Big)\\&=\frac{i}{2\hbar}\Big(\hat{\vec{p}}[\hat{\vec{p}},\hat{x}_\alpha]_-+[\hat{\vec{p}},\hat{x}_\alpha]_-\hat{\vec{p}}\Big)\\&=\frac{i}{2\hbar}\Big(\hat{\vec{p}}\frac{\hbar}{i}\vec{e}_\alpha+\frac{\hbar}{i}\vec{e}_\alpha\hat{\vec{p}}\Big)\\&=\vec{e}_\alpha\hat{\vec{p}}=\hat{p}_\alpha\\\Rightarrow\quad\hat{p}_\alpha&=\frac{i}{2\hbar}[\hat{\vec{p}}^2,\hat{x}_\alpha]_-\end{aligned} \tag{L.71}$$

With $\hat{H}_0=\frac{\hat{\vec{p}}^2}{2m_e}+V(\vec{r})$, I obtain

$$\hat{p}_\alpha=\frac{im_e}{\hbar}[\hat{H}_0,\hat{x}_\alpha]_- \tag{L.72}$$

The interaction is

$$\begin{aligned}\hat{W}(t)&=-\frac{\hbar q}{m_e}\vec{e}_A\vec{k}\frac{1}{2}\Big(A_0\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}-A_0^*\mathrm{e}^{-i(\vec{k}\vec{r}-\omega t)}\Big)g(t)\\&\quad+\frac{q^2}{2m_e}\frac{1}{4}\Big[\Big(\underbrace{A_0^2\mathrm{e}^{2i(\vec{k}\vec{r}-\omega t)}+(A_0^*)^2\mathrm{e}^{-2i(\vec{k}\vec{r}-\omega t)}}_{\text{overtones}}+|A_0|^2\Big)g^2(t)\\&=-\frac{i\hbar q}{m_e}\vec{e}_A\vec{k}\,\mathrm{Im}\Big(A_0\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}\Big)g(t)\\&\quad+\frac{q^2}{2m_e}\frac{1}{4}\Big[|A_0|^2+\underbrace{2\mathrm{Re}\Big(A_0^2\mathrm{e}^{2i(\vec{k}\vec{r}-\omega t)}\Big)}_{\text{overtones}}\Big]g^2(t)\end{aligned} \tag{L.73}$$

Now, I use $\vec{e}_A = \frac{i}{\hbar}[\vec{p}, (\vec{e}_A\vec{r})]_-$ and obtain

$$\hat{W}(t) = -\frac{i\hbar q}{m_e}\underbrace{\frac{i}{\hbar}[\vec{p}, (\vec{e}_A\vec{r})]_-}_{\vec{e}_A}\vec{k}\,\mathrm{Im}\Big(A_0 \mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}\Big)g(t)$$
$$+\frac{q^2}{2m_e}\frac{1}{4}\left[|A_0|^2 + \underbrace{2\mathrm{Re}\Big(A_0^2 \mathrm{e}^{2i(\vec{k}\vec{r}-\omega t)}\Big)}_{\text{overtones}}\right]g^2(t) \tag{L.74}$$

$$\hat{W}(t) = -\frac{\hbar q A_0}{m}\underbrace{\frac{i}{\hbar}[\vec{p}, (\vec{e}_A\vec{r})]_-}_{\vec{e}_A}\vec{q}\mathrm{e}^{i(\vec{q}\vec{r}-\omega t)}g(t) + O(A_0)^2 \tag{L.75}$$

Appendix M

Quantum electrodynamics

M.1 The action of quantum electrodynamics

Quantum electrodynamics is the most simple quantum field theory involving particles and fields, namely electrons, positrons and the electromagnetic field. The “classical action” is the starting point for setting up the quantum field theory.

ACTION OF QUANTUM ELECTRODYNAMICS

$$S = \frac{1}{c} \int d^4x \; \mathbf{\Psi}^\dagger \boldsymbol{\gamma}^0 \left[\boldsymbol{\gamma}^\mu (\underbrace{i\hbar\partial_\mu}_{\hat{p}_\mu} - qA_\mu)c - m_0 c^2 \right] \mathbf{\Psi} - \frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu}$$

Here the wave function $\mathbf{\Psi}$ of the electron is a four-component spinor function, which we write as 4×1 matrix. $F^{\mu\nu}$ is the electric field tensor defined as

$$F^{\mu\nu} \stackrel{\text{def}}{=} \partial^\mu A^\nu - \partial^\nu A^\mu$$

The **Dirac matrices** obey the anticommutation relation

$$[\boldsymbol{\gamma}^\mu, \boldsymbol{\gamma}^\nu]_+ = 2g^{\mu,\nu}$$

where $g^{\mu,\nu}$ is the metric tensor. The Dirac matrices can have the form

$$\boldsymbol{\gamma}^0 = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\boldsymbol{\gamma}^1 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_x \\ -\boldsymbol{\sigma}_x & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{\gamma}^2 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_y \\ -\boldsymbol{\sigma}_y & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{\gamma}^3 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_z \\ -\boldsymbol{\sigma}_z & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

In order to avoid transforming the Dirac matrices under Lorentz transformations, we can alternatively transform the spinors so that

$$\begin{aligned} a'^{\mu} &= \Lambda^{\mu}{}_{\nu} a^{\nu} \\ \boldsymbol{\Psi}'^{\dagger}\boldsymbol{\gamma}^{\mu}\boldsymbol{\Psi}' &= \Lambda^{\mu}{}_{\nu}\boldsymbol{\Psi}^{\dagger}\boldsymbol{\gamma}^{\nu}\boldsymbol{\Psi} \end{aligned}$$

Thus

$$\begin{aligned} \Psi' &= \boldsymbol{S}\Psi \\ \boldsymbol{S}^{\dagger}\boldsymbol{\gamma}^{\mu}\boldsymbol{S} &= \Lambda^{\mu}{}_{\nu}\boldsymbol{\gamma}^{\nu} \end{aligned}$$

M.2 Positrons

Here we want to show how that the last two components of the Dirac spinor describe spin-up and spin-down positrons. We do this by deriving the nonrelativistic limit for $E \approx -m_0c^2$. We begin with Eq. 14.4 from p. 230.

$$\begin{pmatrix} m_0c^2 + q\Phi - E & \sum_i \sigma_i(p_i - qA_i)c \\ \sum_i \sigma_i(p_i - qA_i)c & -m_0c^2 + q\Phi - E \end{pmatrix} \begin{pmatrix} |\phi\rangle \\ |\chi\rangle \end{pmatrix} = 0$$

Let us we introduce a small parameter $\epsilon = E + m_0c^2$ and let us introduce the relativistic mass

$$M(\epsilon, \vec{r}) = m_0 + \frac{-\epsilon + q\Phi}{2c^2}$$

$$\begin{pmatrix} 2m_0c^2 + q\Phi - \epsilon & \sum_i \sigma_i(p_i - qA_i)c \\ \sum_i \sigma_i(p_i - qA_i)c & -\epsilon + q\Phi \end{pmatrix} \begin{pmatrix} |\phi\rangle \\ |\chi\rangle \end{pmatrix} = 0$$

so that we obtain

$$\begin{pmatrix} 2Mc^2 & \sum_i \sigma_i(p_i - qA_i)c \\ \sum_i \sigma_i(p_i - qA_i)c & -\epsilon + q\Phi \end{pmatrix} \begin{pmatrix} |\phi\rangle \\ |\chi\rangle \end{pmatrix} = 0$$

Now we use the first equation to express the electronic components by the positronic ones.

$$|\phi\rangle = -\frac{1}{2Mc^2}\sum_i \sigma_i(p_i - qA_i)c|\chi\rangle$$

and insert it into the second equation

$$\begin{aligned}
0 &= \sum_i \sigma_i(p_i - qA_i)c|\phi\rangle - \epsilon + q\Phi|\chi\rangle \\
&= \left[-\sum_i \sigma_i(p_i - qA_i)\frac{1}{2M}\sum_i \sigma_i(p_i - qA_i) + q\Phi - \epsilon\right]|\chi\rangle \\
&= -\left[\sum_i \sigma_i(p_i - qA_i)\frac{1}{2M}\sum_i \sigma_i(p_i - qA_i) - q\Phi + \epsilon\right]|\chi\rangle \\
&= -\left[\sum_i (p_i - qA_i)\frac{1}{2M}(p_i - qA_i) + i\sum_i \sigma_i\left((\vec{p} - q\vec{A}) \times \frac{1}{2M}(\vec{p} - q\vec{A})\right)_i - q\Phi + \epsilon\right]|\chi\rangle
\end{aligned}$$

Now we are ready to form the non-relativistic limit. As c approaches infinity, the electronic component $|\phi\rangle$ goes to zero. The positronic components fulfil

$$0 = \left[\frac{(\vec{p} - q\vec{A})^2}{2m_0} - \frac{q}{m}\vec{B}\vec{S} - q\Phi + \epsilon\right]|\chi\rangle$$

Let us compare this result with the Pauli equation for the electrons

$$0 = \left[\frac{(\vec{p} - q\vec{A})^2}{2m_0} - \frac{q}{m}\vec{B}\vec{S} + q\Phi - \epsilon\right]|\phi\rangle$$

The equation for the positrons differs from that for the electrons by the sign reversal of the energy and the potential energy. In order to extract the implications for the dynamics, we use this equation to extract a classical Hamilton function. From that we can derive the classical equations of motion using Hamilton's equations.

One open problem is that we still have a two component spinor wave function. We want to avoid to deal with a classical spinor particle, even though such a generalization could be made. In order to obtain two one-component wave functions, we restrict the equation to magnetic fields oriented along the z-axis. In this case, the positronic Pauli equation separates exactly into one equation for spin-up and one for spin-down positrons. nevertheless we keep the vector notation for $\vec{S}\vec{B}$.

$$
\begin{aligned}
H_p(\vec{p},\vec{r},t) &= -\left[\frac{(\vec{p}-q\vec{A})^2}{2m_0}-\frac{q}{m}\vec{B}\vec{S}-q\Phi\right] \\
\dot{p}_i &= -\frac{\partial H}{\partial r_i} = \left[\frac{(\vec{p}-q\vec{A})}{m_0}(-q\partial_i\vec{A})-\frac{q}{m}(\partial_i\vec{B})\vec{S}-q\partial_i\Phi\right] \\
\dot{r}_i &= \frac{\partial H}{\partial p_i} = -\frac{p_i-qA_i}{m_0} \\
m_0\ddot{r}_i &= -\dot{p}_i + q\dot{\vec{r}}\vec{\nabla}A_i + q\partial_t A_i \\
&= -\left[\frac{(\vec{p}-q\vec{A})}{m_0}(-q\partial_i\vec{A})-\frac{q}{m}(\partial_i\vec{B})\vec{S}-q\partial_i\Phi\right] + q\dot{\vec{r}}\vec{\nabla}A_i + q\partial_t A_i \\
&= -q\dot{\vec{r}}\partial_i\vec{A} + \frac{q}{m}(\partial_i\vec{B})\vec{S} + q\partial_i\Phi + q\dot{\vec{r}}\vec{\nabla}A_i + q\partial_t A_i \\
&= -q\underbrace{\left[-\partial_i\Phi-\partial_t A_i\right]}_{E_i} - q\underbrace{\sum_j\left[\dot{r}_j\partial_i A_j-\dot{r}_j\partial_j A_i\right]}_{(\vec{v}\times\vec{B})_i} - \frac{q}{m}(\partial_i\vec{B})\vec{S} \\
m_0\ddot{\vec{r}} &= -q\left[\vec{E}+\vec{v}\times\vec{B}\right] - \frac{q}{m}(\partial_i\vec{B})\vec{S}
\end{aligned}
$$

The first contribution is the Lorentz force acting on a particle with charge $-q$. The second part is the force related to the energy of a particle with magnetic moment $\vec{m}_p = \frac{q}{m}\vec{S}$ in a magnetic field $\vec{B}$. The energy of a magnetic moment $\vec{m}$ in a magnetic field $\vec{B}$ is $E = -\vec{m}\vec{B}$. The magnetic moment of a classical charge distribution is $\vec{m} = g\frac{q}{2m}\vec{L}$, where g is the **g-factor**, and $\vec{L}$ is the angular momentum.

Remember that the charge q of the electron is negative $q = -e$ so that the charge of the positron is $-q = +e$, where e is the elementary charge.

Thus we can identify the Dirac spinor

$$
\begin{aligned}
|\Psi_1\rangle &= |\phi_1\rangle \to e^-_\uparrow \\
|\Psi_2\rangle &= |\phi_2\rangle \to e^-_\downarrow \\
|\Psi_3\rangle &= |\chi_1\rangle \to p_\downarrow \\
|\Psi_4\rangle &= |\chi_2\rangle \to p_\uparrow
\end{aligned}
$$

The positron p_σ behaves exactly like an electron e_σ with the opposite charge.

Mathematical appendices

Appendix N

Method of separation of variables

The method of separation of variables is a method to solve partial differential equations. While it is not applicable only in a subset of systems, it leads to so dramatic simplification that it is usually the first method that is tried.

A simple example

Let us consider a simple example first, namely the one-dimensional wave equation

$$\partial_x^2 \Phi(x,t) = \frac{1}{c^2}\partial_t^2 \Phi(x,t)$$

We start with an product-Ansatz

$$\Phi(x,t) = f(x)g(t)$$

During the calculation we have to verify if the Ansatz leads to a solution. If inconsistencies appear, we need to discard the Ansatz.

By dividing through the wave function we obtain

$$\frac{1}{f(x)}\partial_x^2 f(x) = \frac{1}{c^2}\frac{1}{g(t)}\partial_t^2 g(t)$$

Since the right side does not depend on x, the left-hand side must be a constant. The analogous argument holds for the right-hand side. This allows us to break the problem up into two one-dimensional problems, namely

$$\begin{aligned}
\frac{1}{f(x)}\partial_x^2 f(x) &= C\\
\frac{1}{c^2}\frac{1}{g(t)}\partial_t^2 g(t) &= C
\end{aligned}$$

which are two eigenvalue equations.

$$\begin{aligned}
\partial_x^2 f(x) &= Cf(x)\\
\frac{1}{c^2}\partial_t^2 g(x) &= Cg(x)
\end{aligned}$$

The breakup of the partial differential equation into several differential equations of lower dimensionality is the essence of the method of separation of variables. The appearance of Eigenvalue equations is also a fairly common feature of the method of separation of variables.

The method

The general idea of the method of separation of variables is to break up a partial differential equation into problems of lower dimensionality. It starts with a ansatz of a product wave function.

While I am not sure if the following describes the complete class of problems to which the method of separation of variables can be applied, it is by far the most common one.

The method of separation of variables works always, if the differential operator can be written as a sum of terms, where each only acts on a subset of the coordinates. In that case we can break the partial differential equation into those subsets.

$$D(\vec{\nabla}_x, \vec{\nabla}_y, \vec{x}, \vec{y})\Phi(\vec{x}, \vec{y}) = 0$$

with

$$D(\vec{\nabla}_x, \vec{\nabla}_y, \vec{x}, \vec{y}) = D_1(\vec{\nabla}_x, \vec{x}) + D_2(\vec{\nabla}_y, \vec{y})$$

We choose the Ansatz

$$\Phi(\vec{x}, \vec{y}) = f(\vec{x})g(\vec{y})$$

and obtain

$$\begin{aligned}
\left[D_1(\vec{\nabla}_x, \vec{x}) + D_2(\vec{\nabla}_y, \vec{y})\right] f(\vec{x})g(\vec{y}) &= 0 \\
\frac{1}{f(\vec{x})}D_1(\vec{\nabla}_x, \vec{x})f(\vec{x}) + \frac{1}{g(\vec{y})}D_2(\vec{\nabla}_y, \vec{y})g(\vec{y}) &= 0
\end{aligned}$$

The important step is that $D_1 fg = gD_1 f$ and $D_2 fg = fD_2 g$, which is valid, because the differential operators do not act on the coordinates of the function, that is shifted in front of the expression.

This leads to

$$\begin{aligned}
\frac{1}{f(\vec{x})}D_1(\vec{\nabla}_x, \vec{x})f(\vec{x}) &= C \\
\frac{1}{g(\vec{y})}D_2(\vec{\nabla}_y, \vec{y})f(\vec{y}) &= C
\end{aligned}$$

which is equivalent to two eigenvalue equations for the individual subsets

$$\begin{aligned}
D_1(\vec{\nabla}_x, \vec{x})f\vec{x}) &= -Cf(\vec{x}) \\
D_2(\vec{\nabla}_y, \vec{y})g(\vec{y}) &= Cg(\vec{y})
\end{aligned}$$

Appendix O

Trigonometric functions

The zero's of the cosine function lie at odd integer multiples of $\frac{\pi}{2}$ and those of the sin function lie at even multiples of $\frac{\pi}{2}$.

$$
\begin{aligned}
\cos(x) = 0 \qquad &\Rightarrow \qquad x = \frac{\pi}{2}(1+2k) \\
\sin(x) = 0 \qquad &\Rightarrow \qquad x = \frac{\pi}{2}(2k)
\end{aligned}
$$

with integer k.

The cosinus is an even function with respect to sign-change of the argument and the sinus is an odd function

$$
\begin{aligned}
\cos(x) &= \cos(-x) \\
\sin(x) &= -\sin(x)
\end{aligned}
$$

The formulas involving trigonometric functions appear quite complex and hard to memorize. However, they are extremely simple to derive, if one only considers them as real and imaginary part of the exponential function with imaginary argument.

One needs for this the identity

$$
e^{ix} = \cos(x) + i\sin(x) \tag{O.1}
$$

the elementary rules for the exponential function

$$
\begin{aligned}
e^{x+y} &= e^x e^y \\
e^0 &= 1 \\
\partial_x e^x &= e^x \,.
\end{aligned}
$$

We start with the first formula that we will need frequently:

$$
\begin{aligned}
e^{ix}e^{-ix} &= e^{i0} = 1 \\
1 &= \Big(\cos(x) + i\sin(x)\Big)\Big(\cos(x) - i\sin(x)\Big) = \cos^2(x) + \sin^2(x)
\end{aligned}
$$

Thus we obtain

$$
\cos^2(x) + \sin^2(x) = 1 \tag{O.2}
$$

The addition theorem is then obtained in the following way

$$
\begin{aligned}
e^{i(x+y)} &= e^{ix}e^{iy} \\
\cos(x+y) + i\sin(x+y) &\overset{\text{Eq. } O.1}{=} \Big(\cos(x) + i\sin(x)\Big)\Big(\cos(y) + i\sin(y)\Big) \\
&= \cos(x)\cos(y) - \sin(x)\sin(y) + i\Big(\cos(x)\sin(y) + \sin(x)\cos(y)\Big)
\end{aligned}
$$

By comparing real and imaginary part individually we obtain

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y) \tag{O.3}$$
$$\sin(x+y) = cos(x)\sin(y) + \sin(x)\cos(y) \tag{O.4}$$

Similarly we obtain the square of sinus and cosinus as

$$\begin{aligned}
\left(\mathrm{e}^{ix}\right)^2 &= \mathrm{e}^{2ix} \\
\left(\cos(x) + i\sin(x)\right) &\overset{\text{Eq. } O.1}{=} \cos(2x) + i\sin(2x) \\
\cos^2(x) - \sin^2(x) + 2i\sin(x)\cos(x) &= \cos(2x) + i\sin(2x)
\end{aligned} \tag{O.5}$$

With the help of Eq. O.2 we can extract the square of cosinus and sinus functions from the real part of the above identity.

$$\begin{aligned}
\cos^2(x) &= \frac{1}{2}\underbrace{\left(\cos^2(x) + \sin^2(x)\right)}_{=1} + \frac{1}{2}\underbrace{\left(\cos^2(x) - \sin^2(x)\right)}_{\cos(2x)} \\
&= \frac{1}{2}\left(1 + \cos(2x)\right)
\end{aligned} \tag{O.6}$$

$$\begin{aligned}
\sin^2(x) &= \frac{1}{2}\underbrace{\left(\cos^2(x) + \sin^2(x)\right)}_{=1} - \frac{1}{2}\underbrace{\left(\cos^2(x) - \sin^2(x)\right)}_{\cos(2x)} \\
&= \frac{1}{2}\left(1 - \cos(2x)\right)
\end{aligned} \tag{O.7}$$

The derivatives are obtained as

$$\begin{aligned}
\frac{d\mathrm{e}^{ix}}{dx} &= i\mathrm{e}^{ix} \\
\frac{d\cos(x)}{dx} + i\frac{d\sin(x)}{dx} &= -\sin(x) + i\cos(x)
\end{aligned}$$

Thus we obtain

$$\begin{aligned}
\frac{d\cos(x)}{dx} &= -\sin(x) \\
\frac{d\sin(x)}{dx} &= +\cos(x)
\end{aligned}$$

Appendix P

Gauss' theorem

Gauss theorem says that the divergence of a vector field integrated over a volume is equal to the component of the vector field normal to the surface integrated over the surface of the volume

$$\int_\Omega dr \, \sum_i \nabla_i f_i(r) = \sum_i \int_{\partial\Omega} dA_i \, f_i(r) \tag{P.1}$$

It is the multidimensional analogue of

$$\int_a^b dx \frac{df}{dx} = f(b) - f(a)$$

A variant of Gauss' theorem

$$\int_\Omega dr \nabla_i g(r) = \int_{\partial\Omega} dA_i g(r) \tag{P.2}$$

is obtained by using $f_i = e_i g(r)$, where $e_i = \delta_{i,j}$ is a unity vector pointing into the Cartesian direction j.

$$\begin{aligned}\int_\Omega dr \sum_i \nabla_i f(r) &= \int_\Omega dr \sum_i \nabla_i \delta_{i,j} g(r) = \int_\Omega dr \nabla_j g(r) = \sum_i \int_{\partial\Omega} dA_i f_i(r) \\ &= \sum_i \int_{\partial\Omega} dA_i \delta_{i,j} g(r) = \int_{\partial\Omega} dA_j g(r)\end{aligned}$$

Appendix Q

Fourier transform

Q.1 General transformations

We would like to expand the function in a set of functions $\{\phi_i(x)\}$. This set of functions should be complete so that any function $f(x)$ is can be expressed as

$$f(x) = \sum_i \phi_i(x) F_i$$

We obtain the expansion coefficients f_i in the following way. We pick one function out of the set, multiply the complex conjugate from the left and integrate

$$\int dx\ \phi_j^*(x) f(x) = \sum_i \Big[\int dx\ \phi_j^*(x)\phi_i(x)\Big] F_i\ .$$

We construct an inverse $\boldsymbol{U}$ of the overlap matrix, so that

$$\sum_k U_{i,k} \int dx\ \phi_k^*(x)\phi_j(x) = \delta_{i,j}\ .$$

Note that if functions in the set $\{\phi_i\}$ is overcomplete[1], the matrix $\boldsymbol{U}$ is not uniquely defined and cannot be obtained using the conventional matrix inversion techniques. One way to obtain an inverse is singular value decomposition.

Then we multiply the matrix $\boldsymbol{U}$ from the left and obtain

$$F_i = \sum_j U_{i,j} \int dx\ \phi_j^*(x) f(x)$$

Q.2 Fourier transform in an finite interval

Let us now consider a function $f(x)$ that is defined on the interval $[-\frac{L}{2}, \frac{L}{2}]$. We choose a special set of functions, namely $\phi_n(x) = C\mathrm{e}^{iG_n x}$, where $G_n L = 2\pi n$ and where C is an arbitrary constant.

The overlap matrix is

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} dx\ \phi_n^*(x)\phi_m(x) = C^2 \int_{-L/2}^{L/2} dx\ \mathrm{e}^{-i(G_n - G_m)x} = C^2 \int_{-L/2}^{L/2} dx\ \mathrm{e}^{i\frac{2\pi}{L}(m-n)x} = C^2 L\delta_{n,m}$$

[1]A set of functions is over complete, if one function out of the set can be represented as a linear superposition of the others.

Being diagonal, the overlap matrix can be inverted easily so that

$$F_i = \frac{1}{C^2L}\int dx\; C\mathrm{e}^{-iG_nx}f(x) = \frac{1}{CL}\int dx\; \mathrm{e}^{-iG_nx}f(x)$$

Forward an backward transform are therefore

$$f(x) = C\sum_n \mathrm{e}^{iG_nx}F_n \tag{Q.1}$$

$$F_n = \frac{1}{CL}\int_{-L/1}^{L/2} dx\; \mathrm{e}^{-iG_nx}f(x) \tag{Q.2}$$

with $G_nL = 2\pi n$

The choice of the parameter C is arbitrary. The following choices are common.

- A good choice is $C = 1$, so that

$$f(x) = \sum_n \mathrm{e}^{iG_nx}F_n$$

$$F_n = \frac{1}{L}\int_{-L/1}^{L/2} dx\; \mathrm{e}^{-iG_nx}f(x)$$

- Often the symmetric form is preferred.

$$f(x) = \sqrt{\frac{1}{L}}\sum_n \mathrm{e}^{iG_nx}F_n$$

$$F_n = \sqrt{\frac{1}{L}}\int_{-L/1}^{L/2} dx\; \mathrm{e}^{-iG_nx}f(x)$$

Q.3 Fourier transform on an infinite interval

Let us now consider the limit $L \to \infty$. We introduce the spacing $\Delta = \frac{2\pi}{L}$ between reciprocal lattice points. Since this spacing vanishes in the limit of a large interval, we can convert the sum Eq. Q.1 into an integral $\Delta\sum_n \to \int dG$ of the continuous variable G.

$$f(x) \overset{\text{Eq. }Q.1}{=} \frac{C}{\Delta}\int dG\; \mathrm{e}^{iGx}F(G) = \frac{CL}{2\pi}\int dG\; \mathrm{e}^{iGx}F(G)$$

$$F(G) \overset{\text{Eq. }Q.2}{=} \frac{1}{CL}\int_{-L/2}^{L/2} dx\; \mathrm{e}^{-iGx}f(x)$$

The limit can be performed if we keep $\beta \overset{\text{def}}{=} CL$ constant, while letting the interval L go to infinity.

$$f(x) = \beta\int \frac{dG}{2\pi}\; \mathrm{e}^{iGx}F(G)$$

$$F(G) = \frac{1}{\beta}\int dx\; \mathrm{e}^{-iGx}f(x)$$

Different conventions are in use:

- A symmetric form can be obtained by equating the pre factors for the forward and the backward transformation, that is $\beta = \sqrt{2\pi}$.

$$f(x) = \frac{1}{\sqrt{2\pi}}\int dG\; \mathrm{e}^{iGx}F(G)$$

$$F(G) = \frac{1}{\sqrt{2\pi}}\int dx\; \mathrm{e}^{-iGx}f(x)$$

- Personally, I prefer the non-symmetric version with $\beta = 2\pi$, which I am using in the present lecture notes.

$$f(x) = \int \frac{dG}{2\pi}\, \mathrm{e}^{iGx} F(G) \tag{Q.3}$$
$$F(G) = \int dx\, \mathrm{e}^{-iGx} f(x) \tag{Q.4}$$

Q.4 Table of Fourier transforms

$f(x) = \int \frac{dG}{2\pi}\, \mathrm{e}^{iGx} F(G)$	$F(G) = \int dx\, \mathrm{e}^{-iGx} f(x)$
$\delta(x)$	1
e^{iG_0x}	$2\pi\delta(G - G_0)$
$y(x - x_0)$	$y(G)\mathrm{e}^{-iGx_0}$
$\int dx' g(x - x')y(x)$	$g(G)y(G)$

Q.5 Fourier transform of Dirac's delta-function

Let us work out the Fourier transform of the delta function.

The delta function is defined by the equation

$$f(x_0) = \int dx\, \delta(x - x_0) f(x) \;, \tag{Q.5}$$

which must hold for all differentiable functions $f(x)$.

The Fourier transform of the δ-function is a a plane wave in reciprocal space, which is obtained by using $f(x) = \mathrm{e}^{-iGx}$ in Eq. Q.5.

$$\int dx\, \delta(x - x_0)\mathrm{e}^{-iGx} = \mathrm{e}^{-iGx_0} \tag{Q.6}$$

If the delta function is located at the origin, i.e. $x_0 = 0$, the Fourier transform is simply a constant.

By comparing Eq. Q.6 with the equation for the Fourier transform Eq. Q.3 we identify $F(G) = \mathrm{e}^{-iGx_0}$ with the Fourier transform of the delta function. Inserting this into the backward transform Eq. Q.3, we obtain the representation of the delta function as integral over reciprocal space

$$\delta(x - x_0) = \int \frac{dG}{2\pi}\, \mathrm{e}^{iG(x - x_0)}$$

Q.5.1 Fourier transform of the norm

Here, I show that the integral of the absolute squares of a function in the time and the frequency domain are related.

$$\begin{aligned}
\int_{-\infty}^{\infty} dx\, f^*(x)f(x) &\stackrel{\text{Eq. Q.3}}{=} \int \frac{dG}{2\pi} \int \frac{dG'}{2\pi}\, F^*(G')F(G) \lim_{X\to\infty} \int_{-X}^{X} dx\, \mathrm{e}^{i(G-G')x} \\
&= \int \frac{dG}{2\pi} \int \frac{dG'}{2\pi}\, F^*(G')F(G) \lim_{X\to\infty} \frac{\mathrm{e}^{i(G-G')X} - \mathrm{e}^{-i(G-G')X}}{i(G - G')} \\
&= \int \frac{dG}{2\pi} \int \frac{dG'}{2\pi}\, F^*(G')F(G) \lim_{X\to\infty} 2\frac{\sin((G - G')X)}{G - G'} \\
&\stackrel{\text{Eq. Q.8}}{=} \int \frac{dG}{2\pi} \int \frac{dG'}{2\pi}\, F^*(G')F(G) 2\pi\delta(G - G') \\
&= \int \frac{dG}{2\pi} F^*(G)F(G)
\end{aligned} \tag{Q.7}$$

where we used

$$\lim_{a\to\infty} \frac{\sin(ax)}{x} = \pi\delta(x) \ . \tag{Q.8}$$

Eq. Q.8 is shown as follows: The delta function is defined by the result $\int dx\ \delta(x)f(x) = f(0)$ of integrals with "reasonably well-behaved functions" $f(x)$. With the help of the identity[26] $\int_{-\infty}^{\infty} dx\ \frac{\sin(ax)}{x} = \pi$, we find a direct relation to the delta function

$$\int_{-\infty}^{\infty} dx\ f(x) \lim_{a\to\infty} \frac{\sin(ax)}{x} = \int_{-\infty}^{\infty} dy\ \frac{\sin(y)}{y} \lim_{a\to\infty} f\left(\frac{y}{a}\right) = \underbrace{\int_{-\infty}^{\infty} dy\ \frac{\sin(y)}{y}}_{\pi} f(0) = \pi f(0) \ , \tag{Q.9}$$

which establishes Eq. Q.8.

Appendix R

Pauli Matrices

R.1 Definition

All two-by-two matrices can be expressed as a superposition of the unity matrix and three Hermitian two-by-two matrices. These matrices are called Pauli matrices.

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ; \quad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} ; \quad \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{R.1}$$

They play an important role for the description of the electron spin, for any two-state system and for the quantum field theory of Fermions.

R.2 Product table

In order to become familiar with Pauli matrices, let us now investigate the commutator and anti-commutator relations. We use the multiplication table, which is obtained using the two-by-two matrices

$$\begin{aligned} \boldsymbol{\sigma}_j \boldsymbol{\sigma}_{j+1} &= i\boldsymbol{\sigma}_{j+2} \\ \boldsymbol{\sigma}_j \boldsymbol{\sigma}_{j-1} &= -i\boldsymbol{\sigma}_{j-2} \\ \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j &= \mathbf{1} \end{aligned}$$

where the j refers to the indices (x, y, z) with cyclic permutation. That is if $j = y$ then $j + 1 = z$ and $j + 2 = x$. etc. That is,

PRODUCTS OF PAULI MATRICES

$$\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j = \delta_{i,j} \mathbf{1} + i \sum_k \epsilon_{i,j,k} \boldsymbol{\sigma}_k \tag{R.2}$$

The multiplication table has the form ($\boldsymbol{\sigma}_1$ vectical, $\boldsymbol{\sigma}_2$ horizontal)

$\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2$	$\boldsymbol{\sigma}_x$	$\boldsymbol{\sigma}_y$	$\boldsymbol{\sigma}_z$
$\boldsymbol{\sigma}_x$	$\mathbf{1}$	i$\boldsymbol{\sigma}_z$	-i$\boldsymbol{\sigma}_y$
$\boldsymbol{\sigma}_y$	-i$\boldsymbol{\sigma}_z$	$\mathbf{1}$	i$\boldsymbol{\sigma}_x$
$\boldsymbol{\sigma}_z$	i$\boldsymbol{\sigma}_y$	-i$\boldsymbol{\sigma}_x$	$\mathbf{1}$

where $\boldsymbol{\sigma}_1$ refers to the rows, and $\boldsymbol{\sigma}_2$ to the columns.

The commutator table is obtained as the antisymmetric part of the multiplication table

$[\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2]_-$	$\boldsymbol{\sigma}_x$	$\boldsymbol{\sigma}_y$	$\boldsymbol{\sigma}_z$
$\boldsymbol{\sigma}_x$	$\mathbf{0}$	$2i\boldsymbol{\sigma}_z$	$-2i\boldsymbol{\sigma}_y$
$\boldsymbol{\sigma}_y$	$-2i\boldsymbol{\sigma}_z$	$\mathbf{0}$	$2i\boldsymbol{\sigma}_x$
$\boldsymbol{\sigma}_z$	$2i\boldsymbol{\sigma}_y$	$-2i\boldsymbol{\sigma}_x$	$\mathbf{0}$

and the anti-commutator table is the symmetric part

$[\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2]_+$	$\boldsymbol{\sigma}_x$	$\boldsymbol{\sigma}_y$	$\boldsymbol{\sigma}_z$
$\boldsymbol{\sigma}_x$	2	0	0
$\boldsymbol{\sigma}_y$	0	2	0
$\boldsymbol{\sigma}_z$	0	0	2

.

We can derive a very useful formula for Pauli matrices, namely

MAGIC FORMULA

$$(\vec{A}\vec{\boldsymbol{\sigma}})(\vec{B}\vec{\boldsymbol{\sigma}}) = \mathbf{1}(\vec{A}\vec{B}) + i\vec{\boldsymbol{\sigma}}(\vec{A}\times\vec{B}) \tag{R.3}$$

which is proven in the following.

Let us consider two superpositions of Pauli matrices

$$A = \sum_i A_i \boldsymbol{\sigma}_i = \vec{A}\vec{\boldsymbol{\sigma}}$$
$$B = \sum_i B_i \boldsymbol{\sigma}_i = \vec{B}\vec{\boldsymbol{\sigma}}$$

The identity Eq. R.3 can be verified by using the multiplication table.

$$\begin{aligned}
(\vec{A}\vec{\boldsymbol{\sigma}})(\vec{B}\vec{\boldsymbol{\sigma}}) &= \sum_{j,k} A_j B_k \boldsymbol{\sigma}_j \boldsymbol{\sigma}_k \\
&= \sum_j A_j B_j \boldsymbol{\sigma}_j^2 + \sum_{j\neq k} A_j B_k \boldsymbol{\sigma}_j \boldsymbol{\sigma}_k \\
&= \sum_j A_j B_j \mathbf{1} + \sum_{j\neq k} \epsilon_{j,k,l} A_j B_k i\boldsymbol{\sigma}_l \\
&= \vec{A}\vec{B}\mathbf{1} + i\vec{\boldsymbol{\sigma}}(\vec{A}\times\vec{B})
\end{aligned}$$

Hereby $\epsilon_{i,j,k}$ is the completely antisymmetric tensor with $\epsilon_{i,j,k} = 0$ if any pair of the indices are equal. $\epsilon_{i,j,k} = 1$ if (i,j,k) is ascending, that is can be mapped onto the sequence $(1,2,3,1,2)$, and $\epsilon_{i,j,k} = -1$ if (i,j,k) is descending, that is can be mapped onto the sequence $(2,1,3,2,1)$. The vector product can be written as $(\vec{A}\times\vec{B})_i = \sum_{j,k}\epsilon_{i,j,k}A_jB_k$

R.3 Rules for matrices in terms of Pauli matrices

Every complex-valued 2×2 matrix $\boldsymbol{A}$ can be expressed as superposition of the unity and the three Pauli matrices with a complex-valued four-dimensional vector $(a_0, \vec{a})$ with $\vec{a}\in\mathbb{C}^3$ in the form

$$\boldsymbol{A} = a_0\mathbf{1} + \sum_{j=1}^{3} a_j\boldsymbol{\sigma}_j = a_0\mathbf{1} + \vec{a}_j\vec{\boldsymbol{\sigma}} = \begin{pmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3 \end{pmatrix} \tag{R.4}$$

Consequently, we can translate operations on 2×2 matrix into operations on complex-valued four-dimensional vectors. The 2×2 matrix is hermitean if the vector $(a_0,\vec{a})$ is real-valued.

The product of two matrices $\boldsymbol{A}=a_0\mathbf{1}+\vec{a}\vec{\boldsymbol{\sigma}}$ and $\boldsymbol{B}=b_0\mathbf{1}+\vec{b}\vec{\boldsymbol{\sigma}}$ is

$$
\begin{aligned}
\boldsymbol{AB} &\overset{\text{Eq. R.3}}{=} a_0b_0\mathbf{1}+\sum_j(a_0b_j+b_0a_j)\boldsymbol{\sigma}_j+\sum_{m,n}a_mb_n\underbrace{\boldsymbol{\sigma}_m\boldsymbol{\sigma}_n}_{\delta_{m,n}\mathbf{1}+i\sum_j\epsilon_{m,n,j}\boldsymbol{\sigma}_j}\\
&= \Big(a_0b_0+\vec{a}\vec{b}\Big)\mathbf{1}+\Big(a_0\vec{b}+b_0\vec{a}+i\vec{a}\times\vec{b}\Big)\vec{\boldsymbol{\sigma}}
\end{aligned}
\tag{R.5}
$$

$$\boldsymbol{A}^\dagger=a_0^*\mathbf{1}+\vec{a}^*\vec{\boldsymbol{\sigma}} \tag{R.6}$$

$$\mathrm{Tr}[\boldsymbol{A}]=2a_0 \tag{R.7}$$

$$\det[\boldsymbol{A}]=a_0^2-\vec{a}^2 \quad \text{for complex } a_0,\vec{a} \tag{R.8}$$

$$\boldsymbol{A}^{-1}=\frac{1}{a_0^2-\vec{a}^2}\Big(a_0\mathbf{1}-\vec{a}\,\vec{\boldsymbol{\sigma}}\Big) \tag{R.9}$$

$$\lambda_\pm=a_0\pm\sqrt{\vec{a}^2} \quad \text{eigenvalues for complex } a_0,\vec{a} \tag{R.10}$$

A unitary matrix $\boldsymbol{U}$ has the form

$$\boldsymbol{U}^{-1}=\boldsymbol{U}^\dagger\Rightarrow\boldsymbol{U}=\frac{\mathrm{e}^{i\varphi}}{\sqrt{1+\vec{u}^2}}\Big(\mathbf{1}+i\vec{u}\vec{\boldsymbol{\sigma}}\Big) \quad \text{with } \varphi,\vec{u}\in\mathbb{R} \tag{R.11}$$

The eigenvalue problem of a, potentially non-hermitean, matrix has the form

$$\boldsymbol{AU}=\boldsymbol{U\lambda} \tag{R.12}$$

where $\boldsymbol{\lambda}$

$$\boldsymbol{\lambda}=a_0\mathbf{1}+\sqrt{\vec{a}^2}\boldsymbol{\sigma}_3 \tag{R.13}$$

is a diagonal matrix holding the, potentially complex-valued, eigenvalues

$$\lambda_\pm=a_0\pm\sqrt{\vec{a}^2} \quad \text{eigenvalues for complex-valued } a_0,\vec{a} \tag{R.14}$$

The right eigenvectors are expressed in terms of the matrix $\boldsymbol{U}=u_0\mathbf{1}+\vec{u}\vec{\boldsymbol{\sigma}}$, which is only unitary of the matrix $\boldsymbol{A}$ is hermitean. ($\vec{e}_z\overset{\text{def}}{=}(0,0,1)$)

$$
\begin{aligned}
0&=\boldsymbol{AU}-\boldsymbol{U\lambda}=\Big(a_0\mathbf{1}+\vec{a}\vec{\boldsymbol{\sigma}}\Big)\boldsymbol{U}-\boldsymbol{U}\Big(a_0\mathbf{1}+\sqrt{\vec{a}^2}\boldsymbol{\sigma}_3\Big)\\
0&=\frac{\vec{a}}{\sqrt{\vec{a}^2}}\vec{\boldsymbol{\sigma}}\boldsymbol{U}-\boldsymbol{U}\boldsymbol{\sigma}_3=\frac{1}{\sqrt{\vec{a}^2}}\Big(\vec{a}\vec{\boldsymbol{\sigma}}\big(u_0\mathbf{1}+\vec{u}\vec{\boldsymbol{\sigma}}\big)-\big(u_0\mathbf{1}+\vec{u}\vec{\boldsymbol{\sigma}}\big)\sqrt{\vec{a}^2}\vec{e}_z\vec{\boldsymbol{\sigma}}\Big)\\
&=\frac{1}{\sqrt{\vec{a}^2}}\Big(u_0\vec{a}\vec{\boldsymbol{\sigma}}+\vec{a}\vec{u}\mathbf{1}+i(\vec{a}\times\vec{u})\vec{\boldsymbol{\sigma}}-u_0\sqrt{\vec{a}^2}\vec{e}_z\vec{\boldsymbol{\sigma}}-\sqrt{\vec{a}^2}(\vec{u}\vec{e}_z)\mathbf{1}-\sqrt{\vec{a}^2}i(\vec{u}\times\vec{e}_z)\vec{\boldsymbol{\sigma}}\Big)\\
&=\Big(\frac{\vec{a}}{\sqrt{\vec{a}^2}}\vec{u}-\vec{u}\vec{e}_z\Big)\mathbf{1}+\Big(u_0\frac{\vec{a}}{\sqrt{\vec{a}^2}}+i\frac{\vec{a}}{\sqrt{\vec{a}^2}}\times\vec{u}-u_0\vec{e}_z-i\vec{u}\times\vec{e}_z\Big)\vec{\boldsymbol{\sigma}}\\
&=\Big(\frac{\vec{a}}{\sqrt{\vec{a}^2}}-\vec{e}_z\Big)\vec{u}\mathbf{1}+\Big(u_0\Big(\frac{\vec{a}}{\sqrt{\vec{a}^2}}-\vec{e}_z\Big)+i\Big(\frac{\vec{a}}{\sqrt{\vec{a}^2}}+\vec{e}_z\Big)\times\vec{u}\Big)\vec{\boldsymbol{\sigma}}
\end{aligned}
\tag{R.15}
$$

`Editor: This derivation is not complete!`

`Editor: Because` $\det A=a_0^2-\vec{a}^2$`, the invariance of the scalar product in Minkowski space can be translated into the invariance of the determinant of the corresponding hermitean` 2×2 `matrices. Whether this has any physical significance is not clear to me.`

R.4 Diagonalize a two-state Hamiltonian using Pauli matrices

The technique below is frequently used for two-states systems. Instead of working with 2-by-2 matrices, one deals with a three dimensional vectors (plus a constant, which then leads to a four-dimensional vector). Here, I only demonstrate how one can diagonalize a Hamiltonian.

The two-state Hamiltonian can be expressed by the unity and the three Pauli matrices $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$, $\boldsymbol{\sigma}_z$,

$$\begin{aligned}\boldsymbol{H} &= \underbrace{\frac{\bar\epsilon_1+\bar\epsilon_2}{2}}_{a_0}\mathbf{1} + \underbrace{\frac{\bar\epsilon_1-\bar\epsilon_2}{2}}_{a_3}\mathbf{1}\,\boldsymbol{\sigma}_z + \underbrace{\mathrm{Re}[t]}_{a_1}\boldsymbol{\sigma}_x \underbrace{-\mathrm{Im}[t]}_{a_2}\boldsymbol{\sigma}_y = a_0\mathbf{1} + a_1\boldsymbol{\sigma}_x + a_2\boldsymbol{\sigma}_y + a_3\boldsymbol{\sigma}_z \\ &= a_0\mathbf{1} + \vec a\cdot\vec{\boldsymbol{\sigma}}\end{aligned} \tag{R.16}$$

with the four-dimensional vector $(a_0, a_1, a_2, a_3) = (a_0, \vec a)$ having the components

$$a_0 \stackrel{\text{def}}{=} \frac{\bar\epsilon_1+\bar\epsilon_2}{2} \quad\text{and}\quad a_1 \stackrel{\text{def}}{=} \mathrm{Re}[t] \quad\text{and}\quad a_2 \stackrel{\text{def}}{=} -\mathrm{Im}[t] \quad\text{and}\quad a_3 \stackrel{\text{def}}{=} \frac{\bar\epsilon_1-\bar\epsilon_2}{2} \tag{R.17}$$

The four dimensional vector reminds a little on the Minkowsky space in the theory of relativity. (I have not yet looked into this analogy any further.)

The characteristic equation and the resulting eigenvalues are

$$\begin{aligned}\det[\boldsymbol{H}-\epsilon\mathbf{1}] &= \det[(a_0-\epsilon)\mathbf{1} + \vec a\vec{\boldsymbol{\sigma}}] = (a_0-\epsilon+a_3)(a_0-\epsilon-a_3) - (a_1-ia_2)(a_1+ia_2) \\ &= (a_0-\epsilon)^2 - \underbrace{a_1^2 - a_2^2 - a_3^2}_{\vec a^2} \\ \Rightarrow \qquad \epsilon_\pm &= a_0 \pm |\vec a|\end{aligned} \tag{R.18}$$

The equation for the eigenvectors is

$$\begin{aligned}(\boldsymbol{H}-\epsilon\mathbf{1})\vec c &= 0 \\ \Big(a_0\mathbf{1} + \vec a\vec{\boldsymbol{\sigma}} - (a_0 \pm |\vec a|)\mathbf{1}\Big)\vec c_\pm &= 0 \\ \Big(\vec e\vec{\boldsymbol{\sigma}} \mp \mathbf{1}\Big)\vec c_\pm &= 0 \quad\text{with}\quad \vec e \stackrel{\text{def}}{=} \frac{\vec a}{|\vec a|}\end{aligned} \tag{R.19}$$

From here on, I worked out the eigenvectors $\vec c_+$ and $\vec c_-$ in a conventional way and combined them into a unitary matrix $\boldsymbol{U}$,

$$\boldsymbol{U} \stackrel{\text{def}}{=} \begin{pmatrix} c_{1,+} & c_{1,-} \\ c_{2,+} & c_{2,-}\end{pmatrix} = \frac{1}{\sqrt{2(1+e_z)}}\Big(\vec e\vec{\boldsymbol{\sigma}} + \boldsymbol{\sigma}_z\Big) \tag{R.20}$$

which diagonalizes the Hamiltonian, i.e. $\boldsymbol{U}^\dagger\boldsymbol{H}\boldsymbol{U} = a_0\mathbf{1} + |\vec a|\boldsymbol{\sigma}_z$.

Rather than deriving this result, I will test below whether this $\boldsymbol{U}$ diagonalizes the Hamiltonian. In these calculations I will use the magic formula Eq. R.3 for the products of Pauli matrices

$$(\vec a\vec{\boldsymbol{\sigma}})(\vec b\vec{\boldsymbol{\sigma}}) \stackrel{\text{Eq. }R.3}{=} \vec a\vec b\mathbf{1} + i(\vec a\times\vec b)\vec{\boldsymbol{\sigma}} \tag{R.21}$$

First, I test whether the matrix $\boldsymbol{U}$ is unitary

$$\boldsymbol{U}^\dagger\boldsymbol{U} \stackrel{\text{Eq. }R.3}{=} \frac{1}{2(1+e_z)}(e_x^2+e_y^2+(e_z+1)^2)\mathbf{1} = \frac{1+2e_z+1}{2(1+e_z)}\mathbf{1} = \mathbf{1} \tag{R.22}$$

which is the case.

Next, I test whether $\boldsymbol{U}$ diagonalizes the Hamiltonian

$$
\begin{aligned}
\boldsymbol{U}^\dagger \boldsymbol{H}\boldsymbol{U} &= \frac{1}{2(1+e_z)}\Big(\vec{e}\vec{\boldsymbol{\sigma}}+\boldsymbol{\sigma}_z\Big)\Big(a_0\mathbf{1}+|\vec{a}|\vec{e}\vec{\sigma}\Big)\Big(\vec{e}\vec{\boldsymbol{\sigma}}+\boldsymbol{\sigma}_z\Big)\\
&= a_0\mathbf{1}+\frac{|\vec{a}|}{2(1+e_z)}\Big(\vec{e}\vec{\boldsymbol{\sigma}}+\boldsymbol{\sigma}_z\Big)\underbrace{\Big(\vec{e}\vec{\sigma}\Big)\Big(\vec{e}\vec{\boldsymbol{\sigma}}+\boldsymbol{\sigma}_z\Big)}_{\mathbf{1}+(\vec{e}\vec{\boldsymbol{\sigma}})\boldsymbol{\sigma}_z}\\
&\overset{\text{Eq. }R.3}{=} a_0\mathbf{1}+\frac{|\vec{a}|}{2(1+e_z)}\Big(\vec{e}\vec{\boldsymbol{\sigma}}+\boldsymbol{\sigma}_z\Big)\Big(\mathbf{1}+e_z\mathbf{1}+ie_y\boldsymbol{\sigma}_x-ie_x\boldsymbol{\sigma}_y\Big)\\
&= a_0\mathbf{1}+\frac{|\vec{a}|}{2(1+e_z)}\left\{(1+e_z)\begin{pmatrix}e_x\\ e_y\\ e_z+1\end{pmatrix}\vec{\boldsymbol{\sigma}}+i\left[\begin{pmatrix}e_x\\ e_y\\ e_z+1\end{pmatrix}\vec{\boldsymbol{\sigma}}\right]\left[\begin{pmatrix}e_y\\ -e_x\\ 0\end{pmatrix}\vec{\boldsymbol{\sigma}}\right]\right\}\\
&\overset{\text{Eq. }R.3}{=} a_0\mathbf{1}+\frac{|\vec{a}|}{2(1+e_z)}\left\{(1+e_z)\begin{pmatrix}e_x\\ e_y\\ e_z+1\end{pmatrix}\vec{\boldsymbol{\sigma}}-\begin{pmatrix}e_x(e_z+1)\\ e_y(e_z+1)\\ -e_x^2-e_y^2\end{pmatrix}\vec{\boldsymbol{\sigma}}\right\}\\
&= a_0\mathbf{1}+\frac{|\vec{a}|}{2(1+e_z)}\left\{\left[\begin{pmatrix}(1+e_z)e_x-e_x(e_z+1)\\ (1+e_z)e_y-e_y(e_z+1)\\ (e_z+1)^2-e_z^2+1\end{pmatrix}\vec{\boldsymbol{\sigma}}\right]\right\}\\
&= a_0\mathbf{1}+\frac{|\vec{a}|}{2(1+e_z)}(2e_z+2)\boldsymbol{\sigma}_z = a_0\mathbf{1}+|\vec{a}|\boldsymbol{\sigma}_z \quad q.e.d.
\end{aligned}
\tag{R.23}
$$

The last result is a diagonal matrix holding the two eigenvalues $a_0 \pm |\vec{a}|$.

Thus, the first eigenvector is $\vec{c}_+ = (U_{1,1}, U_{2,1})$ is the one for the eigenvalue $\epsilon_+ = a_0 + |\vec{a}|$ and the second eigenvector is $\vec{c}_- = (U_{1,2}, U_{2,2})$ is the one for the eigenvalue $\epsilon_+ = a_0 - |\vec{a}|$.

Appendix S

Matrix identities

Editor: This whole chapter is in preparation

S.1 Notation

We introduce the complete orthonormal set of unit vectors

$$\vec{e}_i$$

which have the components

$$(\vec{e})_i = \vec{e}_i\vec{e}_j = \delta_{i,j}$$

S.2 Identities related to the trace

S.2.1 Definition

The trace of a matrix is defined as the sum of its diagonal elements.

$$\mathrm{Tr}[\boldsymbol{A}] \stackrel{\text{def}}{=} \sum_i A_{i,i} = \sum_{i,j} \delta_{i,j}A_{i,j} \tag{S.1}$$

A matrix can be written in the form

$$\boldsymbol{A} = [\vec{a}_1, \vec{a}_2, \ldots]$$

where the vector components $(\vec{a}_i)_j = A_{i,j}$ or as

$$\boldsymbol{A} = [\sum_i \vec{e}_i A_{i,1}, \sum_i \vec{e}_i A_{i,2}, \sum_i \vec{e}_i A_{i,3}, \ldots] \tag{S.2}$$

S.2.2 Invariance under commutation of a product

$$\mathrm{Tr}[\boldsymbol{AB}] = \mathrm{Tr}[\boldsymbol{BA}]$$

Proof:

$$\mathrm{Tr}[\boldsymbol{AB}] = \sum_{i,j} A_{i,j}B_{j,i} = \sum_{i,j} B_{j,i}A_{i,j} = \mathrm{Tr}[\boldsymbol{BA}]$$

S.2.3 Invariance under cyclic permutation

The trace is invariant under cyclic permutation

$$\mathrm{Tr}[\boldsymbol{ABC}] = \mathrm{Tr}[\boldsymbol{CAB}] \tag{S.3}$$

The proof follows directly from the invariance of the trace under commutation of a product of two matrices: One simply considers the product $\boldsymbol{AB}$ as one matrix and $\boldsymbol{C}$ as the other.

S.2.4 Invariance under unitary transformation

The trace is invariant under transformation by a unitary matrix, i.e.

$$\mathrm{Tr}[\boldsymbol{U}^\dagger \boldsymbol{AU}] = Tr[\boldsymbol{A}] \tag{S.4}$$

$$\mathrm{Tr}[\boldsymbol{U}^\dagger \boldsymbol{AU}] = \sum_{i,j,k} U^\dagger_{i,j} A_{j,k} U_{k,i} = \sum_{i,j,k} U_{k,i} U^\dagger_{i,j} A_{j,k} = Tr[\boldsymbol{UU}^\dagger \boldsymbol{A}] \stackrel{\boldsymbol{UU}^\dagger=\boldsymbol{1}}{=} Tr[\boldsymbol{A}]$$

S.3 Identities related to the determinant

S.3.1 Definition

The determinant is defined by the fully antisymmetric tensor ϵ as

$$\det[\boldsymbol{A}] \stackrel{\mathrm{def}}{=} \sum_{i,j,k,\ldots} \epsilon_{i,j,k,\ldots} A_{i,1} A_{j,2} A_{k,3} \cdots \tag{S.5}$$

The fully antisymmetric tensor has the following elements

- $$\epsilon_{1,2,3,\ldots} = 1 \tag{S.6}$$
- The matrix element vanishes if two indices are equal.
- The matrix element changes its sign for every permutation of two indices

In the reverse the fully antisymmetric tensor can be expressed by the determinant by

$$\det\Big[\vec{e}_i, \vec{e}_j, \vec{e}_k, \ldots\Big] \stackrel{\text{Eq. S.5}}{=} \sum_{i',j',k',\ldots} \epsilon_{i',j',k',\ldots} \delta_{i',i} \delta_{j',j} \delta_{k',k} \cdots = \epsilon_{i,j,k,\ldots} \tag{S.7}$$

The determinant can also be written in the following form

$$\det[\boldsymbol{A}] = \sum_{i,j,k,\ldots} \epsilon_{i,j,k,\ldots} A_{i,1} A_{j,2} A_{k,3} \cdots = \sum_{i,j,k,\ldots} \det[\vec{e}_i, \vec{e}_j, \vec{e}_k, \ldots] A_{i,1} A_{j,2} A_{k,3} \cdots \tag{S.8}$$

S.3.2 Product rule

$$\det[\boldsymbol{AB}] = \det[\boldsymbol{A}] \det[\boldsymbol{B}] \tag{S.9}$$

Proof: The proof is taken from `http://algebra.math.ust.hk/determinant/05_proof/lecture3.shtml#produc`
Fix the matrix $\boldsymbol{A}$ and consider

$$f(\boldsymbol{B}) \stackrel{\text{def}}{=} \det[\boldsymbol{AB}]$$

as a function of the matrix $\boldsymbol{B}$

Let us use the notation $\boldsymbol{A} = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n]$ where a matrix is represented by its column vectors.

A matrix product can then be written as

$$\boldsymbol{AB} = [\boldsymbol{A}\vec{b}_1, \boldsymbol{A}\vec{b}_2 \ldots]$$

Now we derive two properties of the function f, that we will need later:

1.

$$\begin{aligned} f([\ldots, \vec{u}+\vec{v}, \ldots]) &= \det[\boldsymbol{A}[\ldots, \vec{u}+\vec{v}, \ldots]) = \det[[\ldots, \boldsymbol{A}(\vec{u}+\vec{v}), \ldots]) = \det[[\ldots, \boldsymbol{A}\vec{u}+\boldsymbol{A}\vec{v}, \ldots]) \\ &\stackrel{??}{=} \det[[\ldots, \boldsymbol{A}\vec{u}, \ldots]] + \det[[\ldots, \boldsymbol{A}\vec{v}, \ldots]] = f([\ldots, \vec{u}, \ldots]) + f([\ldots, \vec{v}, \ldots]) \end{aligned} \tag{S.10}$$

2.

$$\begin{aligned} f[(\ldots, \vec{u}, \ldots, \vec{v}, \ldots]) &= \det[\boldsymbol{A} \cdot [\ldots, \vec{u}, \ldots, \vec{v}, \ldots]] && \text{(S.11)} \\ &= \det[[\ldots, \boldsymbol{A}\vec{u}, \ldots, \boldsymbol{A}\vec{v}, \ldots]] && \text{(S.12)} \\ &= -\det[[\ldots, \boldsymbol{A}\vec{v}, \ldots, \boldsymbol{A}\vec{u}, \ldots]] && \text{(S.13)} \\ &= -\det[\boldsymbol{A}[\ldots, \vec{v}, \ldots, \vec{u}, \ldots]] && \text{(S.14)} \\ &= -f[(\ldots, \vec{v}, \ldots, \vec{u}, \ldots]) && \text{(S.15)} \end{aligned}$$

Thus we can write

$$\begin{aligned} f([B]) &\stackrel{\text{Eq. S.2}}{=} f[\sum_i \vec{e}_i B_{i,1}, \sum_i \vec{e}_i B_{i,2}, \ldots] \\ &\stackrel{\text{Eq. S.10}}{=} \sum_{i,j,k,\ldots} f[\vec{e}_i, \vec{e}_j, \ldots] B_{i,1} B_{j,2} \cdots \end{aligned} \tag{S.16}$$

Now we interchange the unit vectors in f until they are in ascending order. According to Eq. S.15 every interchange causes a sign change. The sign changes can be encoded by the fully antisymmetric tensor ϵ

$$\begin{aligned} f([B]) &\stackrel{\text{Eq. S.16}}{=} \sum_{i,j,k,\ldots} f[\vec{e}_i, \vec{e}_j, \ldots] B_{i,1} B_{j,2} \cdots \\ &\stackrel{\text{Eq. S.15}}{=} \sum_{i,j,k,\ldots} f[\vec{e}_1, \vec{e}_2, \ldots] \epsilon_{i,j,k,\ldots} B_{i,1} B_{j,2} \cdots \\ &= \underbrace{f[\vec{e}_1, \vec{e}_2, \ldots]}_{\det[\boldsymbol{A}\cdot\boldsymbol{1}]} \underbrace{\sum_{i,j,k,\ldots} \epsilon_{i,j,k,\ldots} B_{i,1} B_{j,2} \cdots}_{\det[\boldsymbol{B}]} \\ &= \det[\boldsymbol{A}] \det[\boldsymbol{B}] \end{aligned}$$

Because $f(\boldsymbol{B}) = \det[\boldsymbol{AB}]$ this proves Eq. S.9.

S.3.3 Permutation of a product

Permutation of a product of two square matrices does not change the determinant

$$\det[\boldsymbol{AB}] = \det[\boldsymbol{BA}] \tag{S.17}$$

The proof relies on the product rule Eq. S.9

$$\det[\boldsymbol{AB}] \stackrel{\text{Eq. S.9}}{=} \det[\boldsymbol{A}] \det[\boldsymbol{B}] = \det[\boldsymbol{B}] \det[\boldsymbol{A}] \stackrel{\text{Eq. S.9}}{=} \det[\boldsymbol{BA}]$$

S.3.4 1

$$\det[\mathrm{e}^{c\boldsymbol{A}}] = \mathrm{e}^{c\mathrm{Tr}[\boldsymbol{A}]} \tag{S.18}$$

Proof: Assuming that $\boldsymbol{A}$ can be diagonalized, let

$$\boldsymbol{A} = \boldsymbol{U}^{\dagger}\boldsymbol{a}\boldsymbol{U} \tag{S.19}$$

where $\boldsymbol{a}$ is a diagonal matrix containing the eigenvalues of a, and $\boldsymbol{U}$ is a unitary matrix containing the eigenvectors.

$$\begin{aligned}
\det[\mathrm{e}^{c\boldsymbol{A}}] &= \det\left[\sum_{n=0}^{\infty}\frac{1}{n!}(c\boldsymbol{A})^n\right] \overset{\text{Eq. S.19}}{=} \det\left[\sum_{n=0}^{\infty}\frac{1}{n!}\left(\boldsymbol{U}^{\dagger}c\boldsymbol{a}\boldsymbol{U}\right)^n\right] \\
&\overset{\boldsymbol{U}^{\dagger}\boldsymbol{U}=\mathbf{1}}{=} \det\left[\boldsymbol{U}^{\dagger}\left(\sum_{n=0}^{\infty}\frac{1}{n!}\left(c\boldsymbol{a}\right)^n\right)\boldsymbol{U}\right] \\
&= \det\left[\boldsymbol{U}^{\dagger}\mathrm{e}^{c\boldsymbol{a}}\boldsymbol{U}\right] \\
&\overset{\text{Eq. ??}}{=} \det\left[\boldsymbol{U}\boldsymbol{U}^{\dagger}\mathrm{e}^{c\boldsymbol{a}}\right] \\
&\overset{\boldsymbol{U}^{\dagger}\boldsymbol{U}=\mathbf{1}}{=} \det\left[\mathrm{e}^{c\boldsymbol{a}}\right] \\
&= \mathrm{e}^{ca_{11}}\cdot\mathrm{e}^{ca_{22}}\cdots \\
&= \mathrm{e}^{ca_{11}+ca_{22}+\cdots} \\
&= \mathrm{e}^{c\mathrm{Tr}[\boldsymbol{a}]} \\
&= \mathrm{e}^{c\mathrm{Tr}[\boldsymbol{U}^{\dagger}\boldsymbol{a}\boldsymbol{U}]} \\
&= \mathrm{e}^{c\mathrm{Tr}[\boldsymbol{A}]}
\end{aligned}$$

Appendix T

Special Functions

Quantum mechanical problems often characterize special functions. A good source to find details about them is Abramovitz *Mathematical Tables* and in a more accessible form the appendices of Messiah *Quantum Mechanics*.

T.1 Bessel and Hankel functions

Bessel and Hankel functions play a role as radial part of the Schrödinger equation with a constant potential.

T.2 Hermite Polynomials

Hermite polynomials are used to construct the solutions of the Schrödinger equation for the one-dimensional harmonic oscillator.

T.3 Legendre Polynomials

The Legendre polynomials appear in the θ dependence of spherical harmonics.

T.4 Laguerre Polynomials

(Source: Merzbacher)

Laguerre Polynomials are are used for the solutions of the hydrogen atom.

$$
\begin{aligned}
[x\partial_x^2 + (1-x)\partial_x + n]L_n(x) &= 0 \\
L_n(0) &= n!
\end{aligned}
$$

The solutions are the Laguerre polynomials

$$
L_n(x) = \mathrm{e}^x \frac{d^n}{dx^n}\left[x^n \mathrm{e}^{-x}\right] = (-1+\partial_x)^n x^n
$$

n	$L_n(x)$
0	1
1	$1-x$
2	x^2-2x+2
3	$-x^3+9x^2-18x+6$

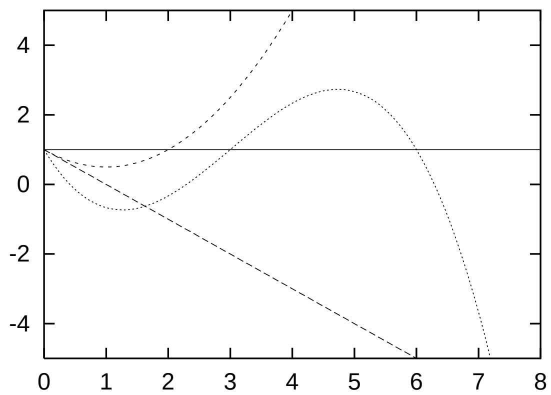


Fig. T.1: Laguerre polynomials rescaled so that the value at the origin is unity

Associated Laguerre Polynomials

$$[x\partial_x^2+(m+1-x)\partial_x+n-m]L_q(x)=0$$

The associated Laguerre polynomials have the form

$$L_n^m(x)=\partial_x^m L_n(x)$$

We start out with the radial Schrödinger equation for the Coulomb potential in Hartree atomic units.

$$\Big[-\frac{1}{2}\partial_r^2+\frac{\ell(\ell+1)}{2r^2}-\frac{Z}{r}-E\Big]r^{\ell+1}\mathrm{e}^{-\lambda r}F(r)$$

Now we use

$$\begin{aligned}
\partial_r r^{\ell+1}\mathrm{e}^{-\lambda r}&=r^{\ell+1}\mathrm{e}^{-\lambda r}\Big(\frac{\ell+1}{r}-\lambda+\partial_r\Big)\\
\partial_r^2 r^{\ell+1}\mathrm{e}^{-\lambda r}&=r^{\ell+1}\mathrm{e}^{-\lambda r}\Big(\frac{\ell+1}{r}-\lambda+\partial_r\Big)^2\\
\Big(\frac{\ell+1}{r}-\lambda+\partial_r\Big)^2&=\partial_r^2-\frac{\ell+1}{r^2}+2\Big(\frac{\ell+1}{r}-\lambda\Big)\partial_r+\Big(\frac{\ell+1}{r}-\lambda\Big)^2\\
&=\partial_r^2+2\Big(\frac{\ell+1}{r}-\lambda\Big)\partial_r-\frac{\ell+1}{r^2}+\frac{(\ell+1)^2}{r^2}-2\lambda\frac{\ell+1}{r}+\lambda^2\\
&=\partial_r^2+2\Big(\frac{\ell+1}{r}-\lambda\Big)\partial_r+\frac{\ell(\ell+1)}{r^2}-2\lambda\frac{\ell+1}{r}+\lambda^2
\end{aligned}$$

Now we insert this to obtain an equation for $F(r)$

$$\Big[\Big(\partial_r^2 + 2\Big(\frac{\ell+1}{r} - \lambda\Big)\partial_r + \frac{\ell(\ell+1)}{r^2} - 2\lambda\frac{\ell+1}{r} + \lambda^2\Big)$$
$$-\frac{\ell(\ell+1)}{r^2} + \frac{2Z}{r} + 2E\Big]F(r) = 0$$
$$\Big[r\partial_r^2 + \Big(2(\ell+1) - 2\lambda r\Big)\partial_r + \Big(2Z - 2\lambda(\ell+1)\Big) + r\Big(2E + \lambda^2\Big)\Big]F(r) = 0$$
$$\Big[x\partial_x^2 + \Big(2(\ell+1) - x\Big)\partial_x + \Big(\frac{Z}{\lambda} - (\ell+1)\Big) + \frac{r}{2\lambda}\Big(2E + \lambda^2\Big)\Big]F(x/(2\lambda) = 0$$

Now we select $\lambda = \sqrt{-2E}$ and obtain

$$\Big[x\partial_x^2 + \Big(2(\ell+1) - x\Big)\partial_x + \Big(\frac{Z}{\lambda} - (\ell+1)\Big)\Big]F(x/(2\lambda) = 0$$

Now we set $p+1 = 2(\ell+1)$ and $q - p = \sqrt{Z^2}{-2E} - \ell - 1$ and obtain ... with $p = 2\ell+1$ and $q = \sqrt{Z^2}{-2E} + \ell$. In order to obtain an integer solution for q we need

$$E = -\frac{Z^2}{2n^2}$$

with $n = q - \ell$ or $q = n + \ell$.

Thus the solution is

$$R(r) = L_{n+\ell}^{2\ell+1}(2\lambda r)r^\ell \mathrm{e}^{-\lambda r}$$
$$\lambda = \sqrt{-2E_n} = \frac{Z}{n}$$

For $n = \sqrt{\frac{Z^2}{-2E}}$ the solutions are the associated Laguerre polynomials $F(x/\sqrt{-2E}) = L_n^\ell(x)$

Appendix U

Principle of least action for fields

Now let us generalize the Formalism of least action to fields: in general we have now a Lagrangian density that depends on the field amplitudes $\phi(x,t)$, their time and spatial derivatives[1] $\partial_t\phi$ and $\partial_x\phi$ respectively and of course on time and spatial coordinates.

$$\mathcal{L}(\phi,\partial_t\phi,\partial_x\phi,x,t)$$

The action is a functional of the field

$$\mathcal{S}([\phi(x,t],x,t)=\int dt\int dx\mathcal{L}(\phi,\partial_t\phi,\partial_x\phi,x,t)$$

The Euler-Lagrange equations are obtained from the variation of $\phi(x,t)$ by $\delta\phi(x,t)$, which yields

$$\begin{aligned}\delta\mathcal{S}([\phi(x,t],x,t)=&\int dt\int dx\Big[\frac{d\mathcal{L}}{d\phi(x,t)}\delta\phi(x,t)+\frac{d\mathcal{L}}{d\partial_t\phi(x,t)}\delta\partial_t\phi(x,t)\\&+\frac{d\mathcal{L}}{d\partial_x\phi(x,t)}\delta\partial_x\phi(x,t)\Big]\\&=\int dt\int dx\delta\phi(x,t)\Big[\frac{d\mathcal{L}}{d\phi(x,t)}-\partial_t\frac{d\mathcal{L}}{d\partial_t\phi(x,t)}-\partial_x\frac{d\mathcal{L}}{d\partial_x\phi(x,t)}\Big]\end{aligned}$$

We have used Gauss theorem, and the requirement that the variation of the field vanishes at the integration boundary, that is at t_i and t_f and at $x=\pm\infty$.

The resulting Euler-Lagrange equations are

$$\partial_t\frac{d\mathcal{L}}{d\partial_t\phi(x,t)}=\frac{d\mathcal{L}}{d\phi(x,t)}-\partial_x\frac{d\mathcal{L}}{d\partial_x\phi(x,t)}$$

We can also derive the Hamilton equations by a Legendre transform of the Lagrangian. We first introduce a field momentum density

$$\Pi(x,t)=\frac{\delta\mathcal{L}}{\delta\partial_t\phi}$$

The field momentum has a different physical meaning than the momentum of mass particles, even though formally they are equivalent.

Next define a Hamilton function

$$H([\Pi(x,t)],[\phi(x,t)],x,t)=\int dx\Big[\Pi(x,t)\partial_t\phi(x,t)-\mathcal{L}\Big]$$

[1] There may be more than one spatial derivatives but we remain here consistent with the one-dimensional description. The generalization is obtained by indexing the spatial coordinates and introducing more of them.

We consider the derivatives of the field amplitudes as independent variables.The derivative of the right hand side with respect to $\partial_t\phi$, vanishes if the Momenta are defined as above.

Let us now consider the derivatives with respect to field amplitudes and momenta.

$$\frac{\delta\mathcal{H}}{\delta\Pi(x,t)} = \partial_t\phi(x,t)$$
$$\frac{\delta\mathcal{H}}{\delta\phi(x,t)} = -\frac{\partial\mathcal{L}}{\partial\phi}$$

Appendix V

ℓ-degeneracy of the hydrogen atom

For the hydrogen atom, the eigenvalues depend only on the main quantum number, but the eigenvalues with the same main angular momenta are degenerate. This degeneracy is not a consequence of rotational symmetry. Rotational symmetry only leads to degenerate eigenvalues with different magnetic quantum numbers m. The ℓ-degeneracy is peculiar[1] for the non-relativistic Schrödinger equation with a Coulomb potential. It is not caused by a geometric symmetry but a so-called **dynamical symmetry**. A symmetry is called dynamical if it originates from the equations of motion and not directly from a space-time symmetry.

I am following a description[53] by M. Jacobi and H. Spahr taken from `http://www.desy.de/~jlouis/Vorlesungen/`

V.1 Laplace-Runge-Lenz Vector

The non-relativistic electron in a hydrogen atom experiences a potential that is analogous to the Kepler problem. The Hamilton operator has the form

$$\hat{H} = \frac{\hat{\vec{p}}^2}{2m} - \frac{\kappa}{|\hat{\vec{r}}|} \tag{V.1}$$

where $\kappa = \frac{Ze^2}{4\pi\epsilon_0}$ for the hydrogen atom.

The classical Kepler problem has a conserved quantity, called the **Laplace-Runge-Lenz vector**

$$\hat{\vec{M}} \stackrel{\text{def}}{=} \frac{1}{2m}\left(\hat{\vec{p}} \times \hat{\vec{L}} - \hat{\vec{L}} \times \hat{\vec{p}}\right) - \frac{\kappa}{|\hat{\vec{r}}|}\hat{\vec{r}} \tag{V.2}$$

This is already the hermitian translation of the classical Runge-Lenz vector.

The classical Laplace-Runge-Lenz vector lies in the plane of the trajectory and points along the principal axis of the ellipse. Conservation of the Runge-Lenz vector implies that the trajectory is a ellipse that is stationary in space. For general spherical potentials there will be a precession of apsides[2] of any orbit. The observation of the precession of the apsides of Mercury has been one supporting experimental evidence for the general theory of relativity, as it was a proof for a deviation of the gravitational force from the pure $1/r$ form.

The Runge-Lenz vector is a constant of motion:

$$[\hat{\vec{M}}, \hat{H}] \stackrel{\text{Eq. V.45}}{=} 0 \tag{V.3}$$

[1] Also the spherical harmonic oscillator in three dimensions exhibits ℓ-degeneracy.

[2] The apsides are the point of smallest and farthest distance from the system's center of gravity. Singular: apsis, plural apsides. The point of closest approach is called periapsis or pericenter. The point of farthest distance is called apoapsis or apocenter. The perihelion is the periapsis of an orbit about the sun.

V.1.1 Commutator relations of Runge-Lenz vector and angular momentum

$$[\hat{L}_i,\hat{L}_j] = i\hbar\sum_k \epsilon_{i,j,k}\hat{L}_k \tag{V.4}$$

$$[\hat{M}_i,\hat{L}_j] \stackrel{\text{Eq. V.53}}{=} i\hbar\sum_k \epsilon_{i,j,k}\hat{M}_k \tag{V.5}$$

$$[\hat{M}_i,\hat{M}_j] \stackrel{\text{Eq. V.54}}{=} i\hbar\frac{-2\hat{H}}{m}\sum_k \epsilon_{i,j,k}\hat{L}_k \tag{V.6}$$

V.1.2 Rescale to obtain a closed algebra

The operators $\hat{\vec{M}}$ and $\hat{\vec{L}}$ would form a closed algebra, if the Hamilton operator would not appear in the last expression Eq. V.6. Because the Hamiltonian commutates with the Runge-Lenz vector, we can define a new operator

$$\hat{\vec{M}}' \stackrel{\text{def}}{=} i\sqrt{\frac{m}{2}}\hat{H}^{-\frac{1}{2}}\hat{\vec{M}} \tag{V.7}$$

This expression is defined because Hamilton operator and Runge-Lenz vector commutate. Thus, *for each cartesian direction,* I can determine a common system of eigenstates $|n\rangle$ with

$$\hat{H}|n\rangle = |n\rangle\epsilon_n$$
$$\hat{M}_i|n\rangle = |n\rangle a_n$$

Then I can write the operator

$$\hat{M}'_i = \sqrt{\frac{m}{2}}\sum_n |n\rangle\frac{a_n}{\sqrt{-\epsilon_n}}\langle n|$$

Note, that the eigensystem n is different for the three components of $\hat{\vec{M}}'$, because the different components of $\hat{\vec{M}}$ do not commutate. The new operator is not hermitian, because it has imaginary eigenvalues for states with positive energy.

Because the Hamiltonian commutates with the Runge-Lenz vector and with the angular momentum, we can easily adapt the commutator relations to the ones for the scaled vector $\vec{M}'$.

$$[\hat{L}_i,\hat{L}_j] = i\hbar\sum_k \epsilon_{i,j,k}\hat{L}_k$$

$$[\hat{M}'_i,\hat{L}_j] \stackrel{\text{Eq. V.5}}{=} i\hbar\sum_k \epsilon_{i,j,k}\hat{M}'_k \tag{V.8}$$

$$[\hat{M}'_i,\hat{M}'_j] \stackrel{\text{Eq. V.6}}{=} i\hbar\sum_k \epsilon_{i,j,k}\hat{L}_k \tag{V.9}$$

V.2 SO(4) symmetry

One introduces the generalized angular momenta

$$\hat{\vec{J}}^{(1)} \stackrel{\text{def}}{=} \frac{1}{2}\left(\hat{\vec{L}}+\hat{\vec{M}}'\right) \tag{V.10}$$

$$\hat{\vec{J}}^{(2)} \stackrel{\text{def}}{=} \frac{1}{2}\left(\hat{\vec{L}}-\hat{\vec{M}}'\right) \tag{V.11}$$

These operators obey the following commutator relations that have been derived in Eq. V.56

$$[\hat{J}_i^{(1)}, \hat{J}_j^{(1)}] = i\hbar \sum_{k=1}^{3} \epsilon_{i,j,k} \hat{J}_k^{(1)} \tag{V.12}$$

$$[\hat{J}_i^{(2)}, \hat{J}_j^{(2)}] = i\hbar \sum_{k=1}^{3} \epsilon_{i,j,k} \hat{J}_k^{(2)} \tag{V.13}$$

$$[\hat{J}_i^{(1)}, \hat{J}_j^{(2)}] = 0 \tag{V.14}$$

This algebra is can be characterized[53] as SO(3)×SO(3), as SU(2)×SU(2), or as SO(4).

SO(d) stands for special orthogonal group in d dimensions. It is the group of real $d \times d$ matrices with determinant +1. These matrices are the rotation matrices in d dimensions. In d dimensions there are $1 + 2 + \ldots + (d-1) = \frac{d(d-1)}{2}$ independent rotation axes corresponding to the independent matrix elements in an antisymmetric matrix in d dimensions. The antisymmetric matrices are the generators of this group.

SU(d) stands for special unitary group. It is the group of complex $d \times d$ matrices with determinant +1.

The SO(4) symmetry corresponds to rotations in a four-dimensional cartesian space. In four dimensions, there are six rotation angles. These rotations do not operate in some physical space.

V.3 ℓ-Degeneracy of the hydrogen atom

In order to make the link to the ℓ-degeneracy, we show that the Hamiltonian of the hydrogen atom can be expressed entirely by the generalized angular momenta $\hat{\vec{J}}^{(1)}$ and $\hat{\vec{J}}^{(2)}$.

$$\begin{aligned}
& & \hat{\vec{M}}^2 &\overset{\text{Eq. V.39}}{=} \frac{2}{m}\hat{H}\left(\hat{\vec{L}}^2 + \hbar^2\right) + \kappa^2 \\
&\Rightarrow & -\frac{m}{2}\hat{H}^{-1}\hat{\vec{M}}^2 + \hat{\vec{L}}^2 + \hbar^2 &= -\frac{m}{2}\kappa^2 \hat{H}^{-1} \\
&\overset{\text{Eq. V.7}}{\Rightarrow} & \hat{\vec{M}}'^2 + \hat{\vec{L}}^2 + \hbar^2 &= -\frac{m}{2}\kappa^2 \hat{H}^{-1} \\
&\overset{Eqs.\ V.10, V.11}{\Rightarrow} & 2\left[\left(\hat{\vec{J}}^{(1)}\right)^2 + \left(\hat{\vec{J}}^{(2)}\right)^2\right] + \hbar^2 &= -\frac{m}{2}\kappa^2 \hat{H}^{-1}
\end{aligned}$$

Thus we can express the Hamiltonian by the squares of the generalized angular momenta $\hat{\vec{J}}^{(1)}$ and $\hat{\vec{J}}^{(2)}$.

$$\Rightarrow \qquad \hat{H} = -\frac{m\kappa^2}{2}\left[2\left(\hat{\vec{J}}^{(1)}\right)^2 + 2\left(\hat{\vec{J}}^{(2)}\right)^2 + \hbar^2\right]^{-1} \tag{V.15}$$

Each of the operators $\hat{\vec{J}}^{(1)}$ and $\hat{\vec{J}}^{(2)}$ obeys an algebra that is identical to the angular momentum algebra, which has been obtained for the normal angular momentum $\hat{\vec{L}}$. Therefore, we obtain the same eigenvalue spectrum, namely

$$\left(\hat{\vec{J}}^{(1)}\right)^2 |j_1, j_2\rangle = |j_1, j_2\rangle \hbar^2 j_1(j_1+1) \tag{V.16}$$

$$\left(\hat{\vec{J}}^{(2)}\right)^2 |j_1, j_2\rangle = |j_1, j_2\rangle \hbar^2 j_2(j_2+1) \tag{V.17}$$

where j_1 and j_2 may be non-negative, half-integer numbers

With Eq. V.55 we can show that the $j_1 = j_2 \stackrel{\text{def}}{=} j$.

$$
\begin{aligned}
0 &\stackrel{\text{Eq. V.55}}{=} \hat{\vec{M}}'\hat{\vec{L}}|j_1, j_2\rangle \stackrel{\text{Eq. V.11}}{=} \left(\hat{\vec{J}}^{(1)} - \hat{\vec{J}}^{(2)}\right)\left(\hat{\vec{J}}^{(1)} + \hat{\vec{J}}^{(2)}\right)|j_1, j_2\rangle \\
&\stackrel{\text{Eq. V.17}}{=} \left(\hat{\vec{J}}^{(1)} - \hat{\vec{J}}^{(2)}\right)|j_1, j_2\rangle\hbar^2\left[j_1(j_1+1) + j_2(j_2+1)\right] \\
&\stackrel{\text{Eq. V.17}}{=} |j_1, j_2\rangle\hbar^2\left[j_1(j_1+1) - j_2(j_2+1)\right]\hbar^2\left[j_1(j_1+1) + j_2(j_2+1)\right] \\
&\Rightarrow \quad j_1 = j_2 \wedge j_1 = -(j_2+1)
\end{aligned}
$$

Since j_1 and j_2 are non-negative, we can conclude that the two parameters are identical.

Now return to Eq. V.15 and determine the eigenvalues

$$
\begin{aligned}
\hat{H}|j\rangle &\stackrel{\text{Eq. V.15}}{=} -\frac{m\kappa^2}{2}\left[2\left(\hat{\vec{J}}^{(1)}\right)^2 + 2\left(\hat{\vec{J}}^{(2)}\right)^2 + \hbar^2\right]^{-1}|j\rangle \\
&= |j\rangle\left[-\frac{m\kappa^2}{2}\left[4\hbar^2 j(j+1) + \hbar^2\right]^{-1}\right] \\
&= |j\rangle\left[-\frac{m\kappa^2}{2\hbar^2}\left[2j+1\right]^{-2}\right]
\end{aligned}
$$

Thus we obtain the eigenvalues

$$
E_n \stackrel{n \stackrel{\text{def}}{=} 2j+1}{=} -\frac{m\kappa^2}{2\hbar^2}\frac{1}{n^2}
$$

This is the eigenvalue spectrum of the hydrogen atom. We can form a basisset

$$
\begin{aligned}
\left(\hat{\vec{J}}^{(1)}\right)^2|j, m_1, m_2\rangle &= |j, m_1, m_2\rangle\hbar^2 j(j+1) \\
\hat{J}_z^{(1)}|j, m_1, m_2\rangle &= |j, m_1, m_2\rangle\hbar m_1 \\
\hat{J}_z^{(2)}|j, m_1, m_2\rangle &= |j, m_1, m_2\rangle\hbar m_2
\end{aligned}
$$

because $\left(\hat{\vec{J}}^{(1)}\right)^2$, $\hat{J}_z^{(1)}$ and $\hat{J}_z^{(1)}$ form with the Hamiltonian a set of mutually commutating operators.

Because the energy depends only on j, there are $(2j+1)^2$ pairs of quantum numbers (m_1, m_2). Hence each eigenvalue has a degeneracy n^2, which is exactly the degeneracy predicted for the hydrogen atom including ℓ-degeneracy.

V.4 Derivation of commutator relations used in this chapter

Here I shall in future include the derivation of the expressions used above: These calculations are not finished. The derivations are from `http://www.theochem.ru.nl/~pwormer/rungelenz.pdf`[54].

This derivation is not completed!!!!

$$
\begin{aligned}
[\hat{p}_i, \hat{L}_j] &= \sum_{m,n}\epsilon_{j,m,n}[\hat{p}_i, \hat{x}_m\hat{p}_n] = \sum_{m,n}\epsilon_{j,m,n}\left(\hat{p}_i\hat{x}_m\hat{p}_n - \hat{x}_m\hat{p}_n\hat{p}_i\right) \\
&= \sum_{m,n}\epsilon_{j,m,n}\left(\frac{\hbar}{i}\delta_{i,m}\hat{p}_n + \underbrace{\hat{x}_m\hat{p}_i\hat{p}_n - \hat{x}_m\hat{p}_n\hat{p}_i}_{=0}\right) = \frac{\hbar}{i}\sum_n\epsilon_{j,i,n}\hat{p}_n \\
&= i\hbar\sum_k\epsilon_{i,j,k}\hat{p}_k
\end{aligned}
\tag{V.18}
$$

$$
\begin{aligned}
[\hat{x}_i, \hat{L}_j] &= \sum_{m,n} \epsilon_{j,m,n}[\hat{x}_i, \hat{x}_m \hat{p}_n] = \sum_{m,n} \epsilon_{j,m,n} \left(\hat{x}_i \hat{x}_m \hat{p}_n - \hat{x}_m \hat{p}_n \hat{x}_i\right) \\
&= \sum_{m,n} \epsilon_{j,m,n} \left(\hat{x}_i \hat{x}_m \hat{p}_n - \hat{x}_m \frac{\hbar}{i}\delta_{n,i} - \hat{x}_m \hat{x}_i \hat{p}_n\right) \\
&= i\hbar \sum_k \epsilon_{i,j,k} \hat{x}_k
\end{aligned} \tag{V.19}
$$

$$
\vec{L}\vec{p} = \sum_{i,j,k} \underbrace{\epsilon_{i,j,k}}_{=-\epsilon_{k,j,i}} x_j \underbrace{p_k p_i}_{p_i p_k} = 0 \tag{V.20}
$$

$$
\vec{L}(\vec{L} \times \vec{p}) = \sum_{i,j,k} \underbrace{\epsilon_{i,j,k}}_{=-\epsilon_{j,i,k}} \underbrace{L_i L_j}_{=L_j L_i} p_k = 0 \tag{V.21}
$$

$$
\begin{aligned}
\vec{L}(\vec{p} \times \vec{L}) &= \sum_{i,j,k} \epsilon_{i,j,k} L_i p_j L_k \\
&\overset{\text{Eq. }V.18}{=} \sum_{i,j,k} \epsilon_{i,j,k} L_i \left(i\hbar \sum_m \epsilon_{j,k,m} p_m\right) + \underbrace{\sum_{i,j,k} \epsilon_{i,j,k} L_i L_k p_j}_{=0} \\
&= i\hbar \sum_{i,m} \underbrace{\left(\sum_{j,k} \epsilon_{i,j,k} \epsilon_{m,j,k}\right)}_{2\delta_{i,m}} L_i p_m = i\hbar \vec{L}\vec{p} \overset{\text{Eq. }V.20}{=} 0
\end{aligned} \tag{V.22}
$$

$$
\begin{aligned}
\vec{L}\frac{\vec{r}}{|\vec{r}|} &= \sum_{i,j,k} \epsilon_{i,j,k} x_j p_k x_i \frac{1}{|\vec{r}|} = \sum_{i,j,k} \epsilon_{i,j,k} x_j \frac{\hbar}{i} \delta_{i,k} \frac{1}{|\vec{r}|} + \sum_{i,j,k} \epsilon_{i,j,k} x_j x_i p_k \frac{1}{|\vec{r}|} \\
&= \frac{\hbar}{i} \sum_{j,j} \underbrace{\epsilon_{i,j,i}}_{=0} x_j \frac{1}{|\vec{r}|} + \sum_{i,j,k} \epsilon_{i,j,k=-\epsilon_{j,i,k}} \underbrace{x_j x_i}_{=x_i x_j} p_k \frac{1}{|\vec{r}|} = 0
\end{aligned} \tag{V.23}
$$

$$
\begin{aligned}
\vec{L}\vec{M} &\overset{\text{Eq. }V.2}{=} \vec{L}\left[\frac{1}{2m}\left(\hat{\vec{p}} \times \hat{\vec{L}} - \hat{\vec{L}} \times \hat{\vec{p}}\right) - \frac{\kappa}{|\hat{\vec{r}}|}\hat{\vec{r}}\right] \\
&= \frac{1}{2m}\left(\underbrace{\vec{L}(\hat{\vec{p}} \times \hat{\vec{L}})}_{=0} - \underbrace{\vec{L}(\hat{\vec{L}} \times \hat{\vec{p}})}_{=0}\right) - \kappa \underbrace{\vec{L}\frac{\hat{\vec{r}}}{|\hat{\vec{r}}|}}_{=0} \overset{Eqs.\ V.21,V.22,V.23}{=} 0
\end{aligned} \tag{V.24}
$$

$$\begin{aligned}
\left(\hat{\vec{L}} \times \hat{\vec{p}}\right)_i &= \sum_{j,k} \epsilon_{i,j,k} \hat{L}_j \hat{p}_k \\
&= \sum_{j,k} \epsilon_{i,j,k} \hat{p}_k \hat{L}_j - \sum_{j,k} \epsilon_{i,j,k} [\hat{p}_k, \hat{L}_j] \\
&\overset{\text{Eq. } V.18}{=} -\sum_{j,k} \epsilon_{i,j,k} \hat{p}_j \hat{L}_k - i\hbar \sum_{j,k} \epsilon_{i,j,k} \sum_m \epsilon_{k,j,m} \hat{p}_m \\
&= -\sum_{j,k} \epsilon_{i,j,k} \hat{p}_j \hat{L}_k - i\hbar \sum_m \underbrace{\left(-\sum_{j,k} \epsilon_{i,j,k} \epsilon_{m,j,k} \right)}_{2\delta_{i,m}} \hat{p}_m \\
&= -\left(\hat{\vec{p}} \times \hat{\vec{L}}\right)_i + 2i\hbar \hat{p}_i
\end{aligned} \tag{V.25}$$

$$\vec{p}(\vec{p} \times \vec{L}) = \sum_{i,j,k} \underbrace{\epsilon_{i,j,k}}_{=-\epsilon_{j,i,k}} \underbrace{p_i p_j}_{=p_j p_i} L_k = 0 \tag{V.26}$$

$$\vec{p}(\vec{L} \times \vec{p}) \overset{\text{Eq. } V.25}{=} \vec{p}\left[-(\vec{p} \times \vec{L}) + 2i\hbar\vec{p}\right] \overset{\text{Eq. } V.26}{=} 2i\hbar\vec{p}^2 \tag{V.27}$$

$$\begin{aligned}
(\vec{p} \times \vec{L})\vec{p} &= \sum_{i,j,k} \epsilon_{i,j,k} p_j L_k p_i = \sum_{i,j,k} \epsilon_{j,k,i} p_j L_k p_i = \vec{p}(\vec{L} \times \vec{p}) \\
&\overset{\text{Eq. } V.27}{=} 2i\hbar\vec{p}^2
\end{aligned} \tag{V.28}$$

V.4.1 Calculation of $\hat{\vec{M}}^2$

Products of $\hat{\vec{p}} \times \hat{\vec{L}}$ and $\hat{\vec{L}} \times \hat{\vec{p}}$

$$\begin{aligned}
(\vec{p} \times \vec{L})^2 &= \sum_{i,j,k,m,n} \epsilon_{i,j,k} \epsilon_{i,m,n} p_j L_k p_m L_n = \sum_{j,k,m,n} \underbrace{\left(\sum_i \epsilon_{i,j,k} \epsilon_{i,m,n} \right)}_{\delta_{j,m}\delta_{k,n} - \delta_{j,n}\delta_{k,m}} p_j L_k p_m L_n \\
&= \sum_{j,k} (p_j L_k p_j L_k - p_j L_k p_k L_j) \\
&= \sum_{j,k} (p_j [L_k, p_j] L_k + p_j p_j L_k L_k - p_j L_k p_k L_j) \\
&\overset{\text{Eq. } V.18}{=} i\hbar \sum_{i,j,k} \underbrace{\epsilon_{k,j,i}}_{=-\epsilon_{k,i,j}} \underbrace{p_j p_i}_{=p_i p_j} L_k + \vec{p}^2 \vec{L}^2 - \sum_j p_j \underbrace{(\vec{L}\vec{p})}_{=0} L_j \\
&\overset{\text{Eq. } V.20}{=} \vec{p}^2 \vec{L}^2
\end{aligned} \tag{V.29}$$

This is the result of Eq. 15 of Wormer[54].

$$
\begin{aligned}
(\vec{p}\times\vec{L})(\vec{L}\times\vec{p}) &\overset{\text{Eq. } V.25}{=} (\vec{p}\times\vec{L})\left[-\hat{\vec{p}}\times\hat{\vec{L}}+2i\hbar\hat{\vec{p}}\right] \\
&\overset{\text{Eq. } V.29}{=} -\vec{p}^2\vec{L}^2+2i\hbar(\vec{p}\times\vec{L})\hat{\vec{p}} \\
&= -\vec{p}^2\vec{L}^2+2i\hbar\sum_{i,j,k}\epsilon_{i,j,k}p_jL_kp_i \\
&= -\vec{p}^2\vec{L}^2+2i\hbar\sum_{i,j,k}\epsilon_{j,k,i}p_jL_kp_i = -\vec{p}^2\vec{L}^2+2i\hbar\underbrace{\vec{p}(\vec{L}\times\vec{p})}_{2i\hbar\vec{p}^2} \qquad (V.30)\\
&\overset{\text{Eq. } V.27}{=} -\vec{p}^2\vec{L}^2-4\hbar^2\vec{p}^2 \qquad (V.31)
\end{aligned}
$$

This is the result of Eq. 18 of Wormer[54].

$$
\begin{aligned}
(\vec{L}\times\vec{p})(\vec{p}\times\vec{L}) &\overset{\text{Eq. } V.25}{=} \left[-\hat{\vec{p}}\times\hat{\vec{L}}+2i\hbar\hat{\vec{p}}\right](\vec{p}\times\vec{L}) \\
&\overset{\text{Eq. } V.29}{=} -\vec{p}^2\vec{L}^2+2i\hbar\underbrace{\vec{p}(\vec{p}\times\vec{L})}_{=0} \\
&\overset{\text{Eq. } V.26}{=} -\vec{p}^2\vec{L}^2 \qquad (V.32)
\end{aligned}
$$

This is the result of Eq. 16 of Wormer[54].

$$
\begin{aligned}
(\vec{L}\times\vec{p})(\vec{L}\times\vec{p}) &\overset{\text{Eq. } V.25}{=} \left[-\hat{\vec{p}}\times\hat{\vec{L}}+2i\hbar\hat{\vec{p}}\right]\left[-\hat{\vec{p}}\times\hat{\vec{L}}+2i\hbar\hat{\vec{p}}\right] \\
&\overset{\text{Eq. } V.29}{=} \vec{p}^2\vec{L}^2-2i\hbar\left[\underbrace{(\vec{p}\times\vec{L})\vec{p}}_{2i\hbar\vec{p}^2}+\underbrace{\vec{p}(\vec{p}\times\vec{L})}_{=0}\right]-4\hbar^2\hat{\vec{p}}^2 \\
&\overset{Eq.\ V.28, V.26}{=} \vec{p}^2\vec{L}^2 \qquad (V.33)
\end{aligned}
$$

This is the result of Eq. 17 of Wormer[54].

We can combine these four results to

$$
\begin{aligned}
\left(\vec{p}\times\vec{L}-\vec{L}\times\vec{p}\right)^2 &= \vec{p}^2\vec{L}^2-(-\vec{p}^2\vec{L}^2-4\hbar^2\vec{p}^2)-(-\vec{p}^2\vec{L}^2)+\vec{p}^2\vec{L}^2 \\
&= 4\vec{p}^2\left(\vec{L}^2+\hbar^2\right) \qquad (V.34)
\end{aligned}
$$

$$
\begin{aligned}
(\vec{L}\times\vec{p})\frac{\vec{r}}{|\vec{r}|} &= \sum_{i,j,k}\epsilon_{i,j,k}L_jp_kx_i\frac{1}{|\vec{r}|} \\
&= \sum_{i,j,k}\epsilon_{i,j,k}L_j\frac{\hbar}{i}\delta_{k,i}\frac{1}{|\vec{r}|}+\sum_{i,j,k}\epsilon_{i,j,k}L_jx_ip_k\frac{1}{|\vec{r}|} \\
&= \sum_{i,j,k}\underbrace{\epsilon_{i,j,i}}_{=0}L_j\frac{\hbar}{i}\frac{1}{|\vec{r}|}+\sum_j L_j\underbrace{\sum_{i,k}\epsilon_{j,k,i}x_ip_k\frac{1}{|\vec{r}|}}_{=-L_j} \\
&= -\frac{\vec{L}^2}{|\vec{r}|} \qquad (V.35)
\end{aligned}
$$

This is the result of Eq. 19 of Wormer[54]. Note that $\hat{\vec{L}}^2$ commutates with any spherical function of $\vec{r}$.

$$
\begin{aligned}
(\vec{p}\times\vec{L})\frac{\vec{r}}{|\vec{r}|} &\overset{\text{Eq. }V.25}{=} -(\vec{L}\times\vec{p})\frac{\vec{r}}{|\vec{r}|}+2i\hbar\vec{p}\frac{\vec{r}}{|\vec{r}|} \\
&\overset{\text{Eq. }V.35}{=} \frac{\vec{L}^2}{|\vec{r}|}+2i\hbar\vec{p}\frac{\vec{r}}{|\vec{r}|}
\end{aligned} \tag{V.36}
$$

This is the result of Eq. 20 of Wormer[54].

$$
\frac{\vec{r}}{|\vec{r}|}(\vec{p}\times\vec{L}) = \frac{1}{|\vec{r}|}\sum_{i,j,k}\epsilon_{i,j,k}x_i p_j L_k = \frac{1}{|\vec{r}|}\vec{L}^2 \tag{V.37}
$$

This is the result of Eq. 21 of Wormer[54].

$$
\begin{aligned}
\frac{\vec{r}}{|\vec{r}|}(\vec{L}\times\vec{p}) &\overset{\text{Eq. }V.25}{=} \frac{\vec{r}}{|\vec{r}|}\left(-\vec{p}\times\vec{L}+2i\hbar\vec{p}\right) = -\frac{\vec{r}}{|\vec{r}|}\left(\vec{p}\times\vec{L}\right)+2i\hbar\frac{\vec{r}}{|\vec{r}|}\vec{p} \\
&\overset{\text{Eq. }V.37}{=} -\frac{\hat{\vec{L}}^2}{|\hat{\vec{r}}|}+2i\hbar\frac{1}{|\vec{r}|}\vec{r}\vec{p}
\end{aligned} \tag{V.38}
$$

This is the result of Eq. 22 of Wormer[54].

$$
\begin{aligned}
\hat{\vec{M}}^2 &\overset{\text{Eq. }V.2}{=} \left[\frac{1}{2m}\left(\hat{\vec{p}}\times\hat{\vec{L}}-\hat{\vec{L}}\times\hat{\vec{p}}\right)-\frac{\kappa}{|\hat{\vec{r}}|}\hat{\vec{r}}\right]^2 \\
&= \frac{1}{4m^2}\underbrace{\left(\hat{\vec{p}}\times\hat{\vec{L}}-\hat{\vec{L}}\times\hat{\vec{p}}\right)^2}_{\text{Eq. }V.34}-\frac{\kappa}{2m}\underbrace{\left(\hat{\vec{p}}\times\hat{\vec{L}}-\hat{\vec{L}}\times\hat{\vec{p}}\right)\left(\frac{\vec{r}}{|\hat{\vec{r}}|}\right)}_{\text{Eq. V.36 and Eq. V.35}}-\frac{\kappa}{2m}\underbrace{\left(\frac{\vec{r}}{|\hat{\vec{r}}|}\right)\left(\hat{\vec{p}}\times\hat{\vec{L}}-\hat{\vec{L}}\times\hat{\vec{p}}\right)}_{\text{Eq. V.37 and Eq. V.38}}+\kappa^2\left(\frac{\vec{r}}{|\hat{\vec{r}}|}\right)^2 \\
&= \frac{1}{4m^2}\left(4\vec{p}^2\left[\vec{L}^2+\hbar^2\right]\right)-\frac{\kappa}{2m}\left(2\frac{\vec{L}^2}{|\vec{r}|}+2i\hbar\vec{p}\frac{\vec{r}}{|\vec{r}|}\right)-\frac{\kappa}{2m}\left(2\frac{\vec{L}^2}{|\vec{r}|}-2i\hbar\frac{\vec{r}}{|\vec{r}|}\vec{p}\right)+\kappa^2 \\
&= \frac{1}{4m^2}\left(4\vec{p}^2\left[\vec{L}^2+\hbar^2\right]\right)-\frac{\kappa}{2m}4\frac{\vec{L}^2}{|\vec{r}|}-\frac{\kappa}{2m}2i\hbar[\vec{p},\frac{\vec{r}}{|\vec{r}|}]+\kappa^2 \\
&= \frac{1}{4m^2}\left(4\vec{p}^2\left[\vec{L}^2+\hbar^2\right]\right)-\frac{\kappa}{2m}4\frac{\vec{L}^2}{|\vec{r}|}-\frac{\kappa}{2m}2\hbar^2\left(\frac{3}{|\vec{r}|}-\frac{1}{|\vec{r}|^3}\vec{r}^2\right)+\kappa^2 \\
&= \frac{2}{m}\left(\frac{\vec{p}^2}{2m}\left[\vec{L}^2+\hbar^2\right]\right)-\frac{\kappa}{2m}4\frac{\vec{L}^2}{|\vec{r}|}-\frac{2\hbar^2}{m}\frac{\kappa}{|\vec{r}|}+\kappa^2 \\
&= \frac{2}{m}\frac{\vec{p}^2}{2m}\left[\vec{L}^2+\hbar^2\right]-\frac{2}{m}\frac{\kappa}{|\vec{r}|}\vec{L}^2-\frac{2}{m}\frac{\kappa}{|\vec{r}|}\hbar^2+\kappa^2 \\
&= \frac{2}{m}\left(\frac{\vec{p}^2}{2m}-\frac{\kappa}{|\vec{r}|}\right)\left[\vec{L}^2+\hbar^2\right]+\kappa^2 \\
&\overset{\text{Eq. }V.1}{=} \frac{2}{m}\hat{H}\left[\vec{L}^2+\hbar^2\right]+\kappa^2
\end{aligned} \tag{V.39}
$$

This is the third statement or Theorem 2 of Wormer[54].

V.4.2 Commutator of $\hat{\vec{M}}$ and $\hat{\vec{H}}$

$$\begin{aligned}
\left([\vec{p}\times\vec{L},\vec{p}^2]\right)_i &= \sum_{i,j,k}\epsilon_{i,j,k}\left(p_jL_kp_m^2 - p_m^2p_jL_k\right)\\
&= \sum_{i,j,k}\epsilon_{i,j,k}\left(p_j(L_kp_m)p_m - p_jp_m(p_mL_k)\right)\\
&= \sum_{i,j,k}\epsilon_{i,j,k}\left(p_j[L_k,p_m]p_m + p_jp_mL_kp_m - p_jp_m[p_m,L_k] - p_jp_mL_kp_m\right)\\
&= \sum_{i,j,k}\epsilon_{i,j,k}\left(-p_j[p_m,L_k]p_m - p_jp_m[p_m,L_k]\right)\\
&\overset{\text{Eq. V.18}}{=} -\sum_{i,j,k}\epsilon_{i,j,k}p_jp_m\left(i\hbar\sum_n\epsilon_{m,k,n}p_n\right)\\
&= -i\hbar\sum_{i,j,n}\underbrace{\left(\sum_k\epsilon_{k,i,j}\epsilon_{k,n,m}\right)}_{\delta_{i,n}\delta_{j,m}-\delta_{i,m}\delta_{j,n}}p_jp_mp_n\\
&= -i\hbar\left(\sum_j p_jp_jp_i - \sum_j\delta_{j,n}p_jp_ip_j\right) = 0
\end{aligned} \tag{V.40}$$

$$\begin{aligned}
\left(\vec{r}\times\vec{L}\right)_i &= \sum_{j,k}\epsilon_{i,j,k}x_jL_k = \sum_{j,k}\epsilon_{i,j,k}x_j\sum_{m,n}\epsilon_{k,m,n}x_mp_n = \sum_{j,n,m}\underbrace{\left(\sum_k\epsilon_{k,i,j}\epsilon_{k,m,n}\right)}_{\delta_{i,m}\delta_{j,n}-\delta_{i,n}\delta_{j,m}}x_jx_mp_n\\
&= \sum_{j,n,m}\delta_{i,m}\delta_{j,n}x_jx_mp_n - \sum_{j,n,m}\delta_{i,n}\delta_{j,m}x_jx_mp_n = \sum_j(x_jx_ip_j - x_jx_jp_i)\\
&= \left(\vec{r}(\vec{r}\vec{p}) - \vec{r}^2\vec{p}\right)_i
\end{aligned} \tag{V.41}$$

This is the result of Eq. 23 of Wormer[54].

$$\begin{aligned}
\left(\vec{L}\times\vec{r}\right)_i &= \sum_{j,k}\epsilon_{i,j,k}L_jx_k = \sum_{j,k}\epsilon_{i,j,k}\sum_{m,n}\epsilon_{j,m,n}x_mp_nx_k = \sum_{k,n,m}\underbrace{\left(\sum_j\epsilon_{j,k,i}\epsilon_{j,m,n}\right)}_{\delta_{k,m}\delta_{i,n}-\delta_{k,n}\delta_{i,m}}x_mp_nx_n\\
&= \sum_{k,n,m}\delta_{k,m}\delta_{i,n}x_mp_nx_k - \sum_{k,n,m}\delta_{k,n}\delta_{i,m}x_mp_nx_k = \sum_k(x_kp_ix_k - x_ip_kx_k)\\
&= \sum_k(-\frac{\hbar}{i}\delta_{k,i}x_k + p_ix_kx_k + \frac{\hbar}{i}\delta_{i,k}x_k - p_kx_kx_i) = \sum_k(p_ix_kx_k - p_kx_kx_i)\\
&= \left(\vec{p}\vec{r}^2 - (\vec{p}\vec{r})\vec{r}\right)_i
\end{aligned} \tag{V.42}$$

This is the result of Eq. 24 of Wormer[54].

$$
\begin{aligned}
[p_i, \frac{x_j}{|\vec{r}|}] &= \frac{\hbar}{i}\left(\frac{\delta_{i,j}}{|\vec{r}|} - \frac{x_i x_j}{|\vec{r}|^3}\right) \\
\rightarrow [p^2, \frac{x_j}{|\vec{r}|}] &= \sum_i \left(p_i p_i \frac{x_j}{|\vec{r}|} \underbrace{-p_i \frac{x_j}{|\vec{r}|} p_i + p_i \frac{x_j}{|\vec{r}|} p_i}_{=0} - \frac{x_j}{|\vec{r}|} p_i p_i \right) \\
&= \sum_i \left(p_i \left[p_i, \frac{x_j}{|\vec{r}|}\right]_- + \left[p_i, \frac{x_j}{|\vec{r}|}\right]_- p_i \right) \\
&= \frac{\hbar}{i} \sum_i \left(p_i \left(\frac{\delta_{i,j}}{|\vec{r}|} - \frac{x_i x_j}{|\vec{r}|^3}\right) + \left(\frac{\delta_{i,j}}{|\vec{r}|} - \frac{x_i x_j}{|\vec{r}|^3}\right) p_i \right) \\
&= \frac{\hbar}{i} \left(p_j \vec{r}^2 \frac{1}{|\vec{r}|^3} - (\vec{p}\vec{x}) \frac{x_j}{|\vec{r}|^3} + \frac{1}{|\vec{r}|^3} \vec{r}^2 p_j - \frac{x_j}{|\vec{r}|^3} (\vec{x}\vec{p}) \right) \\
&= \frac{\hbar}{i} \left[\underbrace{\left(p_j \vec{r}^2 - (\vec{p}\vec{x}) x_j \right)}_{=(\vec{L}\times\vec{r})_j} \frac{1}{|\vec{r}|^3} + \frac{1}{|\vec{r}|^3} \underbrace{\left(\vec{r}^2 p_j - x_j (\vec{x}\vec{p}) \right)}_{-(\vec{r}\times\vec{L})_j} \right] \\
&\overset{Eqs.\ V.42, V.41}{=} \frac{\hbar}{i} \left[(\vec{L}\times\vec{r})_j \frac{1}{|\vec{r}|^3} - \frac{1}{|\vec{r}|^3} (\vec{r}\times\vec{L})_j \right] \\
&\overset{[\vec{r},|\vec{r}|^{-3}]=0;[\vec{L},|\vec{r}|^{-3}]=0}{=} \frac{i\hbar}{|\vec{r}|^3} \left(\vec{r}\times\vec{L} - \vec{L}\times\vec{r} \right)_j
\end{aligned}
\tag{V.43}
$$

$$
\begin{aligned}
\left([\frac{1}{r}, \vec{L}\times\vec{p} - \vec{p}\times\vec{L}] \right)_i &= \sum_{j,k} \epsilon_{i,j,k} \left(\frac{1}{|\vec{r}|} (L_j p_k - p_j L_k) - (L_j p_k - p_j L_k) \frac{1}{|\vec{r}|} \right) \\
&\overset{[\vec{L},|\vec{r}|^{-1}]=0}{=} \sum_{j,k} \epsilon_{i,j,k} \left(L_j \frac{1}{|\vec{r}|} p_k - \frac{1}{|\vec{r}|} p_j L_k - L_j p_k \frac{1}{|\vec{r}|} + p_j \frac{1}{|\vec{r}|} L_k \right) \\
&= \sum_{j,k} \epsilon_{i,j,k} \left(-L_j [p_k, \frac{1}{|\vec{r}|}] + [p_j, \frac{1}{|\vec{r}|}] L_k \right) \\
&= \frac{\hbar}{i} \sum_{j,k} \epsilon_{i,j,k} \left(+L_j \frac{x_k}{|\vec{r}|^3} - \frac{x_k}{|\vec{r}|^3} p_j \right) \\
&= \frac{\hbar}{i} \left(+\vec{L}\times\frac{\vec{r}}{|\vec{r}|^3} - \frac{\vec{r}}{|\vec{r}|^3} \times \vec{L} \right)_j \\
&\overset{[\vec{L},|\vec{r}|^{-3}]=0}{=} \frac{\hbar}{i} \frac{1}{|\vec{r}|^3} \left(\vec{L}\times\vec{r} - \vec{r}\times\vec{L} \right)_j
\end{aligned}
\tag{V.44}
$$

This is the result of Lemma 10 of Wormer[54].

$$
\begin{aligned}
[\hat{\vec{M}}, \hat{H}] &\overset{Eq.\ V.1, V.2}{=} \left[\frac{1}{2m}\left(\hat{\vec{p}} \times \hat{\vec{L}} - \hat{\vec{L}} \times \hat{\vec{p}}\right) - \frac{\kappa}{|\hat{\vec{r}}|}\hat{\vec{r}}, \frac{\hat{\vec{p}}^2}{2m} - \frac{\kappa}{|\hat{\vec{r}}|}\right]_- \\
&= \frac{1}{4m^2}\underbrace{\left[\left(\hat{\vec{p}} \times \hat{\vec{L}} - \hat{\vec{L}} \times \hat{\vec{p}}\right), \vec{p}^2\right]_-}_{=0\ (\text{Eq. V.40,Eq. V.25})} - \frac{\kappa}{2m}\underbrace{\left[\left(\hat{\vec{p}} \times \hat{\vec{L}} - \hat{\vec{L}} \times \hat{\vec{p}}\right), \frac{1}{|\vec{r}|}\right]_-}_{\text{Eq. } V.44} \\
&\quad - \frac{\kappa}{2m}\underbrace{\left[\frac{\hat{\vec{r}}}{|\hat{\vec{r}}|}, \vec{p}^2\right]_-}_{\text{Eq. } V.43} + \kappa^2 \underbrace{\left[\frac{\hat{\vec{r}}}{|\hat{\vec{r}}|}, \frac{1}{|\vec{r}|}\right]_-}_{=0} \\
&= -\frac{\kappa}{2m}\frac{\hbar}{i}\frac{1}{|\vec{r}|^3}\left(\vec{L} \times \vec{r} - \vec{r} \times \vec{L}\right) - \frac{\kappa}{2m}\frac{-i\hbar}{|\vec{r}|^3}\left(\vec{r} \times \vec{L} - \vec{L} \times \vec{r}\right) \\
&= 0 \qquad \text{(V.45)}
\end{aligned}
$$

V.4.3 Commutators of $\hat{\vec{M}}$ and $\hat{\vec{L}}$

$$
\begin{aligned}
[(\vec{L} \times \vec{p})_i, (\vec{L} \times \vec{p})_j] &= \sum_{k,l}\sum_{m,n} \epsilon_{i,k,l}\epsilon_{j,m,n}\left(L_k p_l L_m p_n - L_m p_n L_k p_l\right) \qquad \text{(V.46)} \\
&= \sum_{k,l}\sum_{m,n} \epsilon_{i,k,l}\epsilon_{j,m,n}\left(L_k [p_l, L_m] p_n - L_m [p_n, L_k] p_l\right) \qquad \text{(V.47)} \\
&= \sum_{k,l}\sum_{m,n} \epsilon_{i,k,l}\epsilon_{j,m,n}\left(L_k\left(i\hbar\sum_o \epsilon_{l,m,o}p_o\right)p_n - L_m\left(i\hbar\sum_o \epsilon_{n,k,o}p_o\right)p_l\right) \\
&= i\hbar\sum_{k,l,m,n,o}\left(\epsilon_{i,k,l}\epsilon_{j,m,n}\epsilon_{l,m,o}L_k p_o p_n - \epsilon_{i,k,l}\epsilon_{j,m,n}\epsilon_{n,k,o}L_m p_o p_l\right) \qquad \text{(V.48)} \\
&= i\hbar\sum_{k,l,o}\left(\sum_n \epsilon_{i,k,l}\left(\sum_m \epsilon_{m,n,j}\epsilon_{m,o,l}\right)L_k p_o p_n - \sum_m \epsilon_{i,k,l}\left(\sum_n \epsilon_{j,m,n}\epsilon_{n,k,o}\right)L_m p_o p_l\right) \\
&= i\hbar\sum_{k,l,o}\left(\sum_n \epsilon_{i,k,l}\left(\delta_{n,o}\delta_{j,l} - \delta_{n,l}\delta_{j,o}\right)L_k p_o p_n - \sum_m \epsilon_{i,k,l}\left(\delta_{j,k}\delta_{m,o} - \delta_{j,o}\delta_{m,k}\right)L_m p_o p_l\right) \\
&= i\hbar\sum_{k,l,o}\Big(\sum_n \epsilon_{i,k,l}\delta_{n,o}\delta_{j,l}L_k p_o p_n - \sum_n \epsilon_{i,k,l}\delta_{n,l}\delta_{j,o}L_k p_o p_n \\
&\quad - \sum_m \epsilon_{i,k,l}\delta_{j,k}\delta_{m,o}L_m p_o p_l + \sum_m \epsilon_{i,k,l}\delta_{j,o}\delta_{m,k}L_m p_o p_l\Big) \\
&= i\hbar\left(\sum_{k,n}\epsilon_{i,k,j}L_k p_n p_n - \underbrace{\sum_{k,n}\epsilon_{i,k,n}L_k p_j p_n}_{(A)} - \sum_{l,m}\epsilon_{i,j,l}L_m p_m p_l + \underbrace{\sum_{l,m}\epsilon_{i,m,l}L_m p_j p_l}_{(A)}\right) \\
&= -i\hbar\left(\sum_k \epsilon_{i,j,k}L_k\vec{p}^2 - (\vec{L}\vec{p})\sum_k \epsilon_{i,j,k}p_k\right) \\
&= -i\hbar\sum_k \epsilon_{i,j,k}\left(L_k\vec{p}^2 - (\vec{L}\vec{p})p_k\right) \qquad \text{(V.49)}
\end{aligned}
$$

This differs from the corresponding result of Lemma 12 of Wormer[54].

$$
\begin{aligned}
[(\vec{p}\times\vec{L})_i,(\vec{p}\times\vec{L})_j] &= \sum_{k,l,m,n}\epsilon_{i,k,l}\epsilon_{j,m,n}(p_kL_lp_mL_n - p_mL_np_kL_l)\\
&= \sum_{k,l,m,n}\epsilon_{i,k,l}\epsilon_{j,m,n}(p_k[L_l,p_m]L_n + p_kp_mL_lL_n - p_m[L_n,p_k]L_l - p_mp_kL_nL_l)\\
&\overset{\text{Eq. }V.18}{=} \sum_{k,l,m,n}\epsilon_{i,k,l}\epsilon_{j,m,n}\left(p_k\left(-i\hbar\sum_o\epsilon_{m,l,o}\hat{p}_o\right)L_n - p_m\left(-i\hbar\sum_p\epsilon_{k,n,p}\hat{p}_p\right)L_l\right)\\
&= -i\hbar\sum_{k,l,m,n}\epsilon_{i,k,l}\epsilon_{j,m,n}\left(\sum_o\epsilon_{m,l,o}p_k\hat{p}_oL_n - \sum_p\epsilon_{k,n,p}p_m\hat{p}_pL_l\right)\\
&= -i\hbar\left(\sum_{k,n,o}\left(\sum_l\epsilon_{i,k,l}\sum_m\epsilon_{m,n,j}\epsilon_{m,l,o}\right)p_k\hat{p}_oL_n - \sum_{l,m,o}\left(\sum_k\epsilon_{i,k,l}\sum_n\epsilon_{n,j,m}\epsilon_{n,o,k}\right)p_m\hat{p}_oL_l\right)\\
&= -i\hbar\Big(\sum_{k,n,o}\left(\sum_l\epsilon_{i,k,l}\left(\delta_{n,l}\delta_{j,o}-\delta_{n,o}\delta_{j,l}\right)\right)p_k\hat{p}_oL_n\\
&\quad -\sum_{l,m,o}\left(\sum_k\epsilon_{i,k,l}\left(\delta_{j,o}\delta_{m,k}-\delta_{j,k}\delta_{m,o}\right)\right)p_m\hat{p}_oL_l\Big)\\
&= -i\hbar\left(\sum_{k,n}\epsilon_{i,k,n}p_k\hat{p}_jL_n - \sum_{k,n}\epsilon_{i,k,j}p_k\hat{p}_nL_n - \sum_{l,m}\epsilon_{i,m,l}p_m\hat{p}_jL_l + \sum_{l,m}\epsilon_{i,j,l}p_m\hat{p}_mL_l\right)\\
&= -i\hbar\left(\underbrace{\sum_{k,n}\epsilon_{i,k,n}p_k\hat{p}_jL_n}_{(A)} - \sum_{k,n}\epsilon_{i,k,j}p_k\hat{p}_nL_n - \underbrace{\sum_{k,n}\epsilon_{i,k,n}p_k\hat{p}_jL_n}_{(A)} + \sum_{k,n}\epsilon_{i,j,n}p_k\hat{p}_kL_n\right)\\
&= -i\hbar\left(+\sum_{k,n}\epsilon_{i,j,k}p_k(\hat{\vec{p}}\hat{\vec{L}}) + \vec{p}^2\sum_n\epsilon_{i,j,n}L_n\right) \qquad \text{(V.50)}
\end{aligned}
$$

$$
\begin{aligned}
[(\vec{p}\times\vec{L})_i,p_j] &= \sum_{m,n}\epsilon_{i,m,n}(p_mL_np_j - p_jp_mL_n) = \sum_{m,n}\epsilon_{i,m,n}(p_m[L_n,p_j] + \underbrace{p_mp_jL_n - p_jp_mL_n}_{=0})\\
&\overset{\text{Eq. }V.18}{=} \sum_{m,n}\epsilon_{i,m,n}p_m\left(-i\hbar\sum_k\epsilon_{j,n,k}\hat{p}_k\right)\\
&= -i\hbar\sum_{k,m}\left(\sum_n\epsilon_{n,i,m}\epsilon_{n,k,j}\right)\hat{p}_m\hat{p}_k\\
&= -i\hbar\sum_{k,m}\left(\delta_{i,k}\delta_{m,j}-\delta_{i,j}\delta_{m,k}\right)\hat{p}_m\hat{p}_k\\
&= -i\hbar\left(\hat{p}_i\hat{p}_j - \delta_{i,j}\vec{p}^2\right) \qquad \text{(V.51)}
\end{aligned}
$$

$$
\begin{aligned}
[(\vec{p}\times\vec{L})_i,(\vec{L}\times\vec{p})_j] &\overset{\text{Eq. }V.25}{=} -[(\vec{p}\times\vec{L})_i,(\vec{L}\times\vec{p})_j] + 2i\hbar[(\vec{p}\times\vec{L})_i,p_j]\\
&\overset{Eqs.\ V.50,V.51}{=} i\hbar\left(+\sum_{k,n}\epsilon_{i,j,k}p_k(\hat{\vec{p}}\hat{\vec{L}}) + \vec{p}^2\sum_n\epsilon_{i,j,n}L_n\right) + 2i\hbar\left(-i\hbar\left(\hat{p}_i\hat{p}_j - \delta_{i,j}\vec{p}^2\right)\right)\\
&= i\hbar\sum_{k,n}\epsilon_{i,j,k}p_k(\hat{\vec{p}}\hat{\vec{L}}) + i\hbar\vec{p}^2\sum_n\epsilon_{i,j,n}L_n + 2\hbar^2\left(\hat{p}_i\hat{p}_j - \delta_{i,j}\vec{p}^2\right) \qquad \text{(V.52)}
\end{aligned}
$$

These equations are still given without derivation.

$$[\hat{M}_i, \hat{L}_j] = i\hbar \sum_k \epsilon_{i,j,k} \hat{M}_k \tag{V.53}$$

$$[\hat{M}_i, \hat{M}_j] = i\hbar \frac{-2\hat{H}}{m} \sum_k \epsilon_{i,j,k} \hat{L}_k \tag{V.54}$$

$$\hat{\vec{M}}' \hat{\vec{L}} = 0 \tag{V.55}$$

V.4.4 Commutators of $\hat{\vec{J}}^{(1)}$ and $\hat{\vec{J}}^{(2)}$

We determine the commutator relations of $\hat{\vec{J}}^{(1)}$ and $\hat{\vec{J}}^{(2)}$. The signs in the definitions of the two operators are encrypted by the variables σ. $\sigma = +1$ for the operator $\hat{\vec{J}}^{(1)}$ and $\sigma = -1$ for the operator $\hat{\vec{J}}^{(2)}$.

$$\begin{aligned}
\left[\frac{1}{2}\left(\hat{L}_i + \sigma_1 \hat{M}'_i\right), \frac{1}{2}\left(\hat{L}_j + \sigma_2 \hat{M}'_j\right)\right] &= \frac{1}{4}\left[\hat{L}_i + \sigma_1 \hat{M}'_i, \hat{L}_j + \sigma_2 \hat{M}'_j\right] \\
&= \frac{1}{4}\left(\left[\hat{L}_i, \hat{L}_j\right] + \sigma_2 \underbrace{\left[\hat{L}_i, \hat{M}'_j\right]}_{-\left[\hat{M}'_j, \hat{L}_i\right]} + \sigma_1 \left[\hat{M}'_i, \hat{L}_j\right] + \sigma_1\sigma_2 \left[\hat{M}'_i, \hat{M}'_j\right]\right) \\
&= \frac{1}{4} i\hbar \sum_k \epsilon_{i,j,k} \left(\hat{L}_k + \sigma_2 \hat{M}'_k + \sigma_1 \hat{M}'_k + \sigma_1\sigma_2 \hat{K}_k\right) \\
&= i\hbar \sum_k \epsilon_{i,j,k} \frac{1}{2}\left(\frac{1+\sigma_1\sigma_2}{2}\hat{L}_k + \frac{\sigma_1+\sigma_2}{2}\hat{M}'_k\right) \\
\Rightarrow \quad & \begin{cases}
\left[\hat{J}_i^{(1)}, \hat{J}_i^{(1)}\right] \overset{\sigma_1=\sigma_2=1}{=} i\hbar \sum_k \epsilon_{i,j,k} \hat{J}_k^{(1)} \\
\left[\hat{J}_i^{(1)}, \hat{J}_i^{(2)}\right] \overset{\sigma_1=-\sigma_2=1}{=} 0 \\
\left[\hat{J}_i^{(2)}, \hat{J}_i^{(2)}\right] \overset{\sigma_1=\sigma_2=-1}{=} i\hbar \sum_k \epsilon_{i,j,k} \hat{J}_k^{(2)}
\end{cases}
\end{aligned} \tag{V.56}$$

Appendix W

A small Dictionary

abelian	abelsch
abbreviation	Abkürzung (of a word)
action	Wirkung
adjoint	adjungiert
angle of refraction	Brechungswinkel
angular frequency	Kreisfrequenz
associative law	Assoziativgesetz
arbitrary	willkürlich, beliebig
bead	glasperle
to blur	verschwimmen
bracket	Klammer
branch	Zweig
canned	aus der Dose
chord	Saite
to commutate	vertauschen, vertauschbar sein.
commutative Law	Kommutativgesetz
continuous	stetig
convolution	Faltung
degeneracy	Entartung
degenerate	entartet
differentiable	differenzierbar
distributive law	Distributivgesetz
dyadic product	dyadisches Produkt,äußeres Produkt
elevated	erhöht, hoch
expectation value	Erwartungswert
gauge symmetry	Eichsymmetrie
gem stone	Edelstein
Hamiltonian	Hamilton operator
hermitian	hermitesch
incident angle	Einfallswinkel
integer	ganze Zahl
intricacy	Kniffligkeit
Lagrangian	Lagrange Funktion
Lattice	Gitter
Lettuce	Salat
ledge	Absatz, sims
mean square deviation	mittlere quadratische Abweichung
momentum	Impuls

normalizable	normalisierbar
periodicity	Periodizität
pitch	Grundton
potential well	Potentialtopf
power	Potenz
power series expansion	Potenzreihenentwicklung
principle of least action	Wirkungsprinzip
precaution	Vorsicht
principal axis	Hauptachse
probability	Wahrscheinlichkeit
q.e.d	quod erat demonstrandum= Was zu zeigen war
radius of convergence	Konvergenzradius
realm	Fachgebiet / Gefilde / Reich
rectifier	Gleichrichter
reference frame	Bezugssystem
refraction	Brechung
sequence	Folge
series	Reihe
set	Menge
slope	Steigung
sophisticated	verfeinert, anspruchsvoll, ausgeklügelt
space	Raum
spatial	räumlich
spring constant	Federkonstante
squared	quadratiert
square root	Wurzel
square variation	quadratische Abweichung
state	Zustand
strain	Dehnung
string	Saite
tension	Spannung (elastisch)
transition probability	Übergangswahrscheinlichkeit
trough	Tal
wave crest	Wellenberg
well	Graben

Appendix X

Greek Alphabet

A	α	alpha	N	ν	nu
B	β	beta	Ξ	ξ	ksi
Γ	γ	gamma	O	$o,$	omicron
Δ	δ	delta	Π	π, ϖ	pi
E	ϵ, ε	epsilon	P	ρ, ϱ	rho
Z	ζ	zeta	Σ	σ, ς	sigma
H	η	eta	T	τ	tau
Θ	θ, ϑ	theta	Υ	υ	upsilon
I	ι	iota	Φ	ϕ, φ	phi
K	κ	kappa	X	χ	chi
Λ	λ	lambda	Ψ	ϕ	phi
M	μ	mu	Ω	ω	omega

Appendix Y

Philosophy of the ΦSX Series

In the ΦSX series, I tried to implement what I learned from the feedback given by the students which attended the courses and that relied on these books as background material.

The course should be **self-contained**. There should not be any statements "as shown easily..." if, this is not true. The reader should not need to rely on the author, but he should be able to convince himself, if what is said is true. I am trying to be as complete as possible in covering all material that is required. The basis is the mathematical knowledge. With few exceptions, the material is also developed in a sequence so that the material covered can be understood entirely from the knowledge covered earlier.

The derivations shall be **explicit**. The novice should be able to step through every single step of the derivation with reasonable effort. An advanced reader should be able to follow every step of the derivations even without paper and pencil.

All **units** are explicit. That is, formulas contain all fundamental variables, which can be inserted in any desirable unit system. Expressions are consistent with the SI system, even though I am quoting some final results in units, that are common in the field.

The equations that enter a specific step of a derivation are noted as **hyperlinks** ontop of the equation sign. The experience is that the novice does not immediately memorize all the material covered and that he is struggling with the math, so that he spends a lot of time finding the rationale behind a certain step. This time is saved by being explicit about it. The danger that the student gets dependent on these indications, is probably minor, as it requires some effort for the advanced reader to look up the assumptions, an effort he can save by memorizing the relevant material.

Important results and equations are highlighted by including them in **boxes**. This should facilitate the preparations for examinations.

Portraits of the key researchers and short biographical notes provide independent associations to the material. A student may not memorize a certain formula directly, but a portrait. From the portrait, he may associate the correct formula. The historical context provides furthermore an independent structure to organize the material.

The two first books are in german (That is the intended native language) in order to not add complications to the novice. After these first books, all material is in English. It is mandatory that the student masters this language. Most of the scientific literature is available only in English. English is currently the language of science, and science is absolutely dependent on international contacts.

I tried to include many graphs and figures. The student shall become used to use all his senses in particular the **visual sense**.

I have slightly modified the selection of the material commonly tought in most courses. Some topics, which I consider of mostly historical relevance I have removed. Others such as the Noether theorem, I have added. Some, like chaos, stochastic processes, etc. I have not added yet.

Appendix Z

About the Author

Prof. Dr. rer. nat Peter E. Blöchl studied physics at Karlsruhe University of Technology in Germany. Subsequently he joined the Max Planck Institutes for Materials Research and for Solid-State Research in Stuttgart, where he developed of electronic-structure methods related to the LMTO method and performed first-principles investigations of interfaces. He received his doctoral degree in 1989 from the University of Stuttgart.

Following his graduation, he joined the renowned T.J. Watson Research Center in Yorktown Heights, NY in the US on a World-Trade Fellowship. In 1990 he accepted an offer from the IBM Zurich Research Laboratory in Ruschlikon, Switzerland, which had just received two Nobel prices in Physics (For the Scanning Tunneling Microscope in 1986 and for the High-Temperature Superconductivity in 1987). He spent the summer term 1995 as visiting professor at the Vienna University of Technology in Austria, from where he was later awarded the habilitation in 1997. In 2000, he left the IBM Research Laboratory after a 10-year period and accepted an offer to be professor for theoretical physics at Clausthal University of Technology in Germany. Since 2003, Prof. Blöchl is member of the Braunschweigische Wissenschaftliche Gesellschaft (Academy of Sciences).

The main thrust of Prof. Blöchl's research is related to ab-initio simulations, that is, parameter-free simulation of materials processes and molecular reactions based on quantum mechanics. He developed the Projector Augmented Wave (PAW) method, one of the most widely used electronic structure methods to date. This work has been cited over 45,000 times.[1] It is among the 100 most cited scientific papers of all times and disciplines[2] , and it is among the 10 most-cited papers out of more than 500,000 published in the 120-year history of Physical Review. [3] Next to the research related to simulation methodology, his research covers a wide area from biochemistry, solid state chemistry to solid state physics and materials science. Prof. Blöchl contributed to 8 Patents and published about 100 research publications, among others in well-known Journals such as "Nature". The work of Prof. Blöchl has been cited over 60,000 times, and he has an H-index of 44.[4]

[1] Researcher ID: B-3448-2012

[2] R. van Noorden, B. Maher and R. Nuzzo, Nature 514, 550 (2014)

[3] Oct. 15, 2014, search in the Physical Review Online Archive with criteria "a-z".

[4] Researcher ID: B-3448-2012

Bibliography

[1] M.Alonso and E.J. Finn, *Quantenphysik*, (Addison Wesley, Bonn, 1988); M.Alonso and E.J. Finn, *Fundamental University Physics, Vol III, Quantum and Statistical Physics* (Addison Wesley, 1968)

[2] A. Messiah and G.M. Temmer, *Quantum Mechanics* Vol. I+II,(North Holland, Amsterdam, 1964).

[3] C.Cohen-Tannoudji and F. Laloe, *Quantenmechanik*, (Walter de Gruyter, Berlin, 1999).

[4] W. Nolting, *Grundkurs Theoretische Physik 5,Quantenmechanik Teil 1+2*, (Verlag Zimmermann-Neufang, Ulmen 1992).

[5] D.J. Griffiths,*Introduction to Quantum mechanics*, (Prentice Hall, New jersey, USA, 1994)

[6] "Relativistic Quantum Mechanics", J..D. Bjorken and S.D. Drell, Int' Ser. in Pure and Applied Physics, (McGraw-Hill, 1964) "Relativistische Quantenmechanik" J.D. Bjorken and S.D. Drell, Bibliographisches Institut, Mannheim.

[7] H. Haken, *Quantenfeldtheorie des Festkörpers*, (Teubner, Stuttgart 1973)

[8] J.D. Jackson, *Classical Electrodynamics*, (John Wiley, New York, 1998)

[9] *Quantities, Units and Symbols in Physical Chemistry* I. Mills et al. (Blackwell Scientific, Oxford 1993).

[10] `http://www.codata.org/databases/databases.html`

[11] M. Abramowitz and I.A. Stegun,*Pocketbook of Mathematical Functions*, (Verlag Harri Deutsch, Thun, 1984)

[12] K. Baumann und R.U. Sexl, *Die Deutungen der Quantentheorie*, (Vieweg) [Überblick Über die Interpretationsprobleme der Quantenmechanik sowie 12 wichtige Originalarbeiten in deutscher Sprache]

[13] J. Audretsch, *verschränkte Welt* (Wiley-VCH) [Allgemeinverständliche, auf Vorträgen beruhende Beiträge verschiedener Autoren.]

[14] R. Omnes,*Understanding Quantum Mechanics* (Princeton University Press)[Populärwissenschaftlich gehaltrenes Buch zur Iterpretation der Quantenmechanik, das auch das Dekohärenzkonzept behandelt, aber im wesentlichen auf den Consistent Theories Zugang abhebt.]

[15] D. Giulini et al. *Decoherence and the Appearance of the Classical World in Quantum theory.*(Springer)[Umfassendes, durchaus technisches Werk zu allen Aspekten der Dekoherenz]

Bibliography

[1] Leonard I. Schiff. Quantum mechanics. International series in pure and applied physics. McGraw-Hill Education (India) Pvt Limited, 1968. ISBN 9780070702431. URL http://books.google.de/books?id=XujUaYBU5oEC.

[2] S. Gasiorowicz. Quantum Physics, 3Rd Ed. Wiley India Pvt. Limited, 2007. ISBN 9788126511174. URL http://books.google.de/books?id=qFtQiVmjWUEC.

[3] E. Merzbacher. Quantum Mechanics. Wiley, 1998 (1st ed: 1961 ?). ISBN 9780471887027. URL http://books.google.de/books?id=6Ja_QgAACAAJ.

[4] P.W. Atkins. Molecular Quantum Mechanics. Oxford University Press, 1983 (1st ed. 1970). ISBN 9780198551805. URL http://books.google.de/books?id=Cq0sAAAAYAAJ.

[5] C. Kiefer. Quantentheorie. [Fischer-Taschenbücher] / Fischer-Taschenbücher. Fischer-Taschenbuch-Verlag, 2004. ISBN 9783596153565. URL http://books.google.de/books?id=UHIrcAAACAAJ.

[6] J.L. Basdevant. Lectures on Quantum Mechanics. Springer, 2007. ISBN 9780387377445. URL http://books.google.de/books?id=-mCI1J2x9C8C.

[7] L. de Broglie. Recherches sur la théorie des quanta. Ann. Phys.(Paris), 3:22, 1925.

[8] C. Davisson and L.H. Germer. Diffraction of electrons by a crystal of nickel. Phys. Rev., 30: 705, 1927.

[9] G.P. Thomson and A. Reid. Difraction of cathode rays by a thin film. Nature, 119:890, 1927.

[10] C. Jönsson. Elektroneninterferenzen an mehreren künstlich hergestellten feinspalten. Zeitschr. Phys., 161:454, 1961.

[11] C. Jönsson. Electron diffraction at multiple slits. Am. J. Phys., 42:4, 1974.

[12] C. Jönsson. Elektroneninterferenzen an mehreren künstlich hergestellten feinspalten. Zeitschr. Phys. A, 161:454, 1961.

[13] A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki, and H. Ezawa. Demonstration of single-electron build-up of an interference pattern. Am. J. Phys., 57:117, 1989.

[14] P. E. Blöchl. ΦSX: Theoretische Physik I, Klassische Mechanik. Blöchl, 2015. URL http://www2.pt.tu-clausthal.de/atp/phisx.html.

[15] E. Schrödinger. Quantisierung als eigenwertproblem (erste mitteilung). Ann. Phys.(Leipzig), 1926.

[16] E. Schrödinger. Quantisierung als eigenwertproblem (zweite mitteilung). Ann. Phys.(Leipzig), 1926.

[17] E. Schrödinger. über das verhältnis der heisenberg-born-jordanschen quantenmechanik zu der meinen. Ann. Phys.(Leipzig), 80:437, 1926.

[18] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. Phys. Rev., 28:1049, 1926.

[19] P.J. Mohr and B.N. Taylor. The fundamental physical constants. Physics Today, 56:guide, 2003.

[20] P.J. Mohr and B.N. Taylor. Codata recommended values of the fundamental physical constants: 1989. Rev. Mod. Phys., 72:351, 2000.

[21] R. Courant. Ein allgemeiner satz zur theorie der eigenfunktionen selbstadjungierter differentialausdrücke,. Nachr. v. d. Ges. d. Wiss.zu Göttingen, page 81, 1923.

[22] J. von Neumann. Allgemeine eigenwerttheorie hermitescher funktionaloperatoren. Math. Annal., 102:49, 1929.

[23] von Neumann. Mathematische Grundlagen der Qantenmechanik. Springer, Berlin, 1932.

[24] John v. Neumann. Die eindeutigkeit der schrödingerschen operatoren. Math. Annal., 104:1432, 1931.

[25] Wolfgang Walter. Einführung in die Theorie der Distributionen. BI-Wiss.Verl., 1994.

[26] I.N. Bronstein and K.A. Semendjajew. Taschenbuch der Mathematik. BSB B.G. eubner Verlagsgesellschaft, Leipzig, 1983.

[27] W. Heisenberg. über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. Zeitschr. Phys., 43:172, 1927.

[28] P.A. Robertson. The uncertainty principle. Phys. Rev., 34:163, 1929.

[29] E. Schrödinger. Zum heisenbergschen unschärfeprinzip. Sitzungsberichte Preuss. Akad. d. Wiss. Phys.-Math. Klasse., 19:296, 1930.

[30] W.H. Zurek. Decoherence and the transition from the quantum to classical. Physics Today, 44:36, 1991.

[31] P. Ehrenfest. Bemerkung b̈er die angenäherte gültigkeit der klassisehen mechanik innerhalb der quantenmechanik. Zeitschr. Phys., 45:455, 1927.

[32] P. E. Blöchl. ΦSX: Theoretische Physik II, Elektrodynamik. Blöchl, 2015. URL http://www2.pt.tu-clausthal.de/atp/phisx.html.

[33] J.R. Chelikowsky and M.L. Cohen. Electronic structure of silicon. Phys. Rev. B, 10:5095, 1974.

[34] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and Brian P. Flannery. Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, 1992. The first edition was still in Fortran, in the second edition , it was called numerical recipes in fortran77 with 2nd volume about fortran 90. The third edition from 2007 is in C++. The second edition is available online under http://www.nrbook.com/a/bookfpdf.php.

[35] J.A. Gaunt. The triplets of helium. Phil. Trans. R. Soc. (London) A, 228:151, 1929.

[36] E.R. Cohen, T. Cvitas, J.G. Frey, B. Holmström, K. Kuchitsu, R. Marquardt, I. Mills, F. Pavese, M. Quack, J. Stohner, H.L. Strauss, M. Takami, and A.J. Thor, editors. Quantities, Units and Sumbols in Physical Chemistry, IUPAC Green Book. IUPAC and RSC Publishing, Cambridge, 3 edition, 2008. URL http://media.iupac.org/publications/books/gbook/IUPAC-GB3-2ndPrinting-Online-22apr2011.pdf.

[37] Peter J. Mohr, Barry N. Tayor, and David B. Newell. Codata recommended values of the physical constants: 2010. Rev. Mod. Phys., 84:1527, 2012.

[38] Peter J. Mohr, David B. Newell, and Barry N. Taylor. Codata recommended values of the fundamental physical constants: 2014. Arxiv, 1507:07956, 2015. URL arxiv.org/abs/1507.07956.

[39] M. Abramowitz and I.A. Stegun, editors. Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, volume 55 of Applied Mathematics Series. National Bureau of Standards, 1964.

[40] F. J. Dyson. The radiation theories of tomonaga, schwinger, and feynman. Phys. Rev., 75:486–502, Feb 1949. doi: 10.1103/PhysRev.75.486. URL https://link.aps.org/doi/10.1103/PhysRev.75.486.

[41] Walter Ritz. über eine neue methode zur lösung gewisser variationsprobleme der mathematischen physik. J. Reine Angew. Math., 135:1, 2009.

[42] G. Wentzel. Eine verallgemeinerung der quantenbedingungen für die zwecke der wellenmechanik. Zeitschr. Phys., 38:518, 1926.

[43] H.A. Kramers. Wellenmechanik und halbzahlige quantisierung. Zeitschr. Phys., 39:828, 1926.

[44] Leon Brillouin. La mécanique ondulatoire de schrödinger une method générale de resolution par approximations successives. Comptes Rendus des Seances de l'Académie des Sciences, 183: 24–26, 1926.

[45] Lord Rayleigh. On the propagation of waves through a stratified medium, with special reference to the question of reflection. Proc. Roy. Soc. A, 86:207, 1912.

[46] Carl D. Anderson. The positive electron. Phys. Rev., 43:491, 1933.

[47] C.G. Darwin. The wave equations of the electron. Proc. Roy. Soc. A, 118:654, 1928.

[48] R. de L. Kronig and W.G. Penney. Quantum mechanics of electrons in crystal lattices. Proc. Roy. Soc. A, 130:499, 1931.

[49] A. Unsöld. Ann. der Phys., 82:355, 1927.

[50] E.P. Wigner and Frederick Seitz. Qualitative analysis of the cohesion in metals. Solid State Physics., 1:97, 1955.

[51] R. Peierls. Zur theorie des diamagnetismus von leitungselektronen. Z. Phys., 80:763, 1933. doi: https://doi.org/10.1007/BF01342591.

[52] Douglas R. Hofstadter. Energy levels and wave functions of bloch electrons in rational and irrational magnetic fields. Phys. Rev. B, 14:2239–2249, Sep 1976. doi: 10.1103/PhysRevB.14.2239. URL https://link.aps.org/doi/10.1103/PhysRevB.14.2239.

[53] Matthias Jacobi and Hendrik Spahr. Das noethertheorem in der quantenmechanik und die so(4)-symmertrie des wasserstoffatoms. Technical report, http://www.desy.de/~jlouis/Vorlesungen/QMII06/vortrag_09.pdf, 2006.

[54] P. Wormer. Aanvulling syllabus "lie theorie voor natuurkundigen" van g.j. heckman (mathematisch instituut ku nijmegen) door p.e.s. wormer (instituut voor theoretische chemie ku nijmegen). Technical report, http://www.theochem.ru.nl/ pwormer/rungelenz.pdf, 2003.

Index