Quantum Theory 1

Typesetting and pictures: Georg Börner and Carsten Timm
Translation: Naomi Herrmann

Version: February 5, 2025

Summer semester 2024 TUD | Dresden University of Technology Institute of Theoretical Physics

Contents

The remainder of the lecture notes has not yet been translated.

Chapter 1

Introduction

To regard the basics of quantum theory—the contents of the lecture Quantum Theory 1—as a central part of the study of physics is no exaggeration. It is the first time we systematically discuss the modern, quantum-theoretical description of processes. For now, we restrict ourselves to mechanical systems, the description of which fundamentally differs from our known classical theories, more specifically differing in the connection of formal objects with reality rather than the mathematical methodology. Quantum theoretical predictions are generally probabilistic in nature, i.e., they are statistical statements rather than the deterministic ones known from classical physics, even if we have a full knowledge of the state of the system. One famous example is in the decay of unstable particles: the decay time of multiple particles is never the same, even if they are of the same kind, and follow a random distribution. This fundamental difference compared to classical physics raises the question of how quantum theory is to be interpreted.

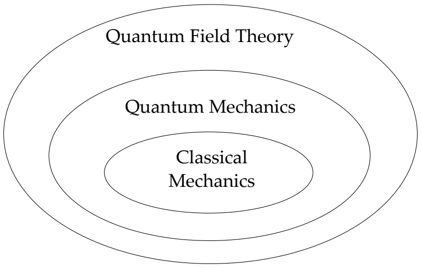
I would like to thank Jan Wüstemann, Bogusz Bujnowski, and Georg Börner for creating various versions of these lecture notes with LaTex as well as creating various depictions. Additionally, I would like to thank all who have brought any mistakes or ambiguities to my attention, most especially Dr. Jürgen Wagner. Finally, I am grateful to Naomi Herrmann for translating these lecture notes into English.

1.1 Goals and Operational Principles of Quantum Theory

Within the broad branches of physics, quantum theory is definitely one of the more abstract fields. We generally have little to no intuition for concepts such as wave equations, operators, and Hilbert spaces. In classical mechanics, the situation is easier since, for example regarding projectile motion, we can catch a ball without having learned a single thing about physics. We have an intuition for understanding these types of motion. However, it is no surprise that we do not posses any intuition for quantum processes since, evolutionarily speaking, we only developed cognitive abilities necessary for survival and interaction with our imminent surroundings. Quantum processes are far removed from any directly observable length and time scales, hence there was no need to react to such processes in the same way as to macroscopic ones, such as predicting the motion of a ball. But rest assured—experience shows it is very much possible for a human to gain an intuition for generally abstract concepts such as quantum theory.

Of course, this doesn't mean that quantum theory is irrelevant for our lives. If we were to live in a universe governed purely by classical physics, nothing would work properly, for example, the sun would not emit light. Or even if the sun did shine there would be no photosynthesis on earth so that the sun light could not be utilized by any animate being. In fact, there would be no earth, much less any life upon it since there would be no stable atoms. Moreover, quantum processes are crucial for technical applications. Without quantum mechanics, one could not understand the functionality of electric components, such as transistors, to name one example. Finally, the entire science of chemistry is applied quantum mechanics.

Quantum mechanics is a theory of particle dynamics. It thus describes the same systems as classical mechanics but it is the more general theory in the sense that quantum mechanics includes classical mechanics as a limiting case. That is at the core of the important *correspondence principle.* Quantum mechanics itself is then embedded into quantum field theory.



In this lecture we will exclusively talk about non-relativistic quantum mechanics. This describes particles with (relative) velocities v which are small compared to the speed of light c. Non-relativistic quantum mechanics is an approximation for $v \ll c$ within relativistic quantum mechanics, which is again included in quantum field theory. Quantum field theory and its application to elementary particles and fields, in sum the standard model of particle physics, does not describe the world entirely. For this, gravity would additionally have to be described, which would require integrating quantum field theory with the (as of yet classical) general theory of relativity. It is still unclear how exactly these two theories could be merged. String theory is a current approach for unifying them.

The goal of quantum mechanics is to find a quantitative description for physical processes. The principle of operation within theoretical physics in general is to formulate a *theory*, a description of a general law. From the more useful theories one can predict the outcome of a given experiment, which in turn allows one to test the theory. As the philosopher of science Sir Karl Popper said, one can never verify theories but one can, in principle, easily disprove (falsify) them. Experiments whose outcomes coincide with a prediction based on a theory can *support* said theory but never prove it. On the other hand, to disprove a theory, a *single* instance is enough that does not coincide with the theoretical prediction. Experiments have the highest authority in physics. So, we can take the stance that quantum mechanics is as it is because extensive experiments support this theory and not an alternative one. Some authors just leave it at that and don't question the interpretation of quantum mechanics, viewing such a question as unproductive. More specifically, this opinion was summed up with the aphorism "shut up and calculate", which apparently goes back to N. D. Mermin, who in fact does not share this view.

In this context, here are a few more points: 1. One should take care to understand what Popper meant by "prove": one cannot prove a theory about the real world *in absolute mathematical sense* but many physical theories are proven in a "legal" sense since they are true according to ones best judgment. The phrase "without reasonable doubt" is perhaps even more apt. 2. An absolute disproof of a theory by an experiment as noted by Popper is also an idealized situation since one cannot be absolutely sure what an experiment actually shows. 3. Many theories are falsified in Popper's sense. As mentioned, non-relativistic quantum mechanics fails at high particle velocities near c. This doesn't mean that the theory is useless or only interesting in historic context. It is actually well understood under which conditions it can render precise predictions. In such cases it wouldn't make sense to use a much more complicated quantum field theory.

Theoretical physics formulates the fundamental laws of nature in the language of mathematics, as this language is best suited for describing quantitative relations. Hence we shall use numerous mathematical methods. Our formulations will mainly use methods from calculus and linear algebra but also from group theory and geometry. Within the lecture, the necessary mathematical means will be developed or reviewed as deemed necessary.

As noted above, the quantum theory that was developed in roughly the last hundred years and has withstood the test of time fundamentally differs in the nature of its predictions from classical theoretical physics. Quantum theory mostly predicts probabilities, as in "the nucleus of a certain isotope decays with a probability of 83% within one second" even if the initial state of the system is fully known. Classical theories are deterministic in their outcome. This fundamental difference calls for a philosophical discussion about the interpretation of quantum theory, which has thus been led since the early days of quantum theory, in parallel to its methodological development and application.

1.2 Outline

The order of topics in this lecture partially follows quantum theory's historical development. This allows us to understand the process of its construction. We aren't simply interested in seeing what modern quantum theory says, rather, we need to understand how it came to be as it is. It's not only our job to preserve existing theories, to apply and to pass them on, it is also our duty to correct and transcend them. The following topics will be expanded on in the course of this lecture:

- The limits of classical physics
- The older quantum theory
- Schrödinger wave mechanics
- Application to one-dimensional systems
- Dirac formalism
- Angular momentum in quantum theory
- Central potentials
- Methods of approximations
- The density operator
- Consequences and interpretations of quantum theory

As an outlook, here are the main topics covered in the lecture Quantum Theory 2:

- Indistinguishable particles and second quantization
- Relativistic quantum theory
- Scattering theory

1.3 Literature

There is a lot of literature about quantum mechanics. Because of the importance of the topic, it is a good idea to own at least one book. Here is an incomplete recommendation list (this information was updated for the English version of these lecture notes):

- W. Nolting, *Theoretical Physics 6: Quantum Mechanics—Basics* (Springer, 2017) and *Theoretical Physics 7: Quantum Mechanics—Methods and Applications* (Springer, 2017): The whole series of textbooks is recommended. Both books on quantum mechanics are relevant for this lecture. Nolting puts a relatively strong emphasis on the training of formalism and correspondingly weaker emphasis on discussions of physics content. He often presents derivations in detail, where other authors only show the results. The presentation is typically clear. The books contain many good exercises with solutions and very helpful control questions.
- A. Messiah, *Quantum Mechanics* (Dover, 2014), single-volume edition corresponding to originally two volumes: Recommended classical textbook. Contains most of the material for the lectures Quantum Theory 1 and 2 and more. The order of presentation differs from the lectures and the distribution of material over the chapters is sometimes confusing. Mathematically rather detailed and containing significantly more discussion than Nolting's books but sometimes a bit wordy. The discussion of the interpretation of quantum mechanics is somewhat outdated. The book contains exercises with solutions.

- D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics* (Cambridge University Press, 2018): U.S. standard textbook for undergraduate quantum mechanics. Didactically and in form and language appealing. Somewhat restricted selection of topics and lower level than Nolting's and Messiah's books and not very systematic. Does not completely cover the present lecture. Dirac formalism is only mentioned in passing. The discussion of those points that are actually covered is very careful and clear. Many exercises without solutions.
- R. L. Liboff, *Introductory Quantum Mechanics* (Addison Wesley, 2003): Another U.S. undergraduate textbook. Somewhat similar to Griffiths' book in its stile but more comprehensive. For example, Liboff's book contains discussions of the foundational experiments, a short introduction to the path integral, and chapters on applications on atomic, molecular, nuclear, and condensed-matter physics. Also covers quite a bit of material from Quantum Theory 2. Discussions are rather shorter than in Griffiths' book and more similar to Nolting's. Contains more modern material, e.g., on the Bell inequality and quantum information (a whole chapter on quantum computing!) than other books but partially outdated remarks on the interpretation of quantum mechanics. Many exercises, partially with solutions.
- S. Weinberg, *Lectures on Quantum Mechanics* (Cambridge University Press, 2015): A graduate-level textbook by Nobel laureate Steven Weinberg. Quantum mechanics is introduced at a very brisk pace. The investigation of central potentials starts on page 32. Many topics are not covered or only mentioned in passing. On the other hand, some topics from Quantum Theory 2 are treated. Does not contain any figures. A strength of Weinberg's book is a rather detailed introduction to the interpretation debate. A few exercises without solutions.
- L. D. Landau and E. M. Lifschitz, *Theoretical physics 3: Quantum mechanics, non-relativistic theory* (Book on Demand, 2018): Part of the well-known classical series of Russian textbooks. By now old fashioned in the selection of topics and presentation. Rather concise and demanding. Intermediate steps of derivations are rarely shown. Might be useful for looking up material if other books do not help. Rather difficult exercises without solutions.
- C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics I: Basic Concepts, Tools, and Applications* (Wiley-VCH, 2019) and *Quantum Mechanics II: Angular Momentum, Spin, and Approximation Methods* (Wiley-VCH, 2019): Very good classical textbook, comprehensive selection of material, which includes essentially all of the lecture Quantum Theory 1 (both volumes are needed) but not the complete lecture Quantum Theory 2. Rather detailed derivations. Contains exercises.
- R. Shankar, *Principles of Quantum Mechanics* (Springer, 2008): U.S. textbook for advanced undergraduate and graduate courses. Many more topics are covered than in Griffiths' book and in this more resembles Messiah's book but in addition contains an introduction to the path integral. The writing stile is informal, even more so than in Griffiths' book, which some readers like while others do not. Contains exercises.

The following books include more modern views on the interpretation of quantum mechanics:

- G. Auletta, *Foundations and Interpretations of Quantum Mechanics* (World Scientific, 2001): Starts with a rather detailed textbook-style introduction to quantum mechanics. The stile of the whole book is more typical for textbooks in mathematics, with definitions and theorems. Contains a very detailed presentation of foundations in information theory, the interpretation debate, and modern topics such as Bell inequality and quantum information, with many references to original literature. The structure of the book is clear but material is sometimes covered in surprising order. Language and layout are somewhat unprofessional (no copy editing?).
- C. Friebe *et al.*, *The Philosophy of Quantum Physics* (Springer, 2018): English translation of a German book from 2015, which is a rather successful attempt to create a textbook for the philosophy of quantum physics. Several well-known German-speaking philosophers of physics have contributed chapters. In view of the topics, the text is rather easy to read. The chapters are relatively independent of one another. The concepts from quantum theory that are meant to be investigated from the philosophic points of view are first introduced carefully. However, also in these introductory sections the book is surprisingly heavy on text as opposed to mathematical material.

- M. Esfeld (editor), *Philosophie der Physik* (Suhrkamp, 2012): Another collection, which contains even more chapters contributed by more authors than the book by Friebe *et al.* Esfeld's book only exists in German, though. It is broader than Friebe's and thus less detailed concerning quantum theory.
- J. Bub, *Interpreting the Quantum Word* (Cambridge University Press, 1997): Not a textbook but a monograph that presents *one* possible interpretation. This interpretation seems to be in the tradition of the idea of de Broglie pilot waves and to integrate modern ideas from information theory. The most relevant part for this lecture is the overview of other interpretations from the point of view of a philosopher of physics not from Germany (in philosophy, national traditions can be important). Contains significantly more mathematical material and also more figures than the German books.
- S. Carroll, *Something Deeply Hidden: Quantum Worlds and the Emergence of Spacetime* (Oneworld Publications, 2021): Carroll is a proponent of the many-worlds interpretation of quantum theory, which he presents to a broad audience in this book. Inspiring and engaging but in view of the target audience not at a deep technical level.

Chapter 2

The Limits of Classical Physics

In the following, we shall discuss which experiments showed that it was necessary to extend physics into the direction of quantum theory. To do so, we must first consider the state of physics during the 19^{th} century.

2.1 The Golden Age of Classical Physics

During the end of the 19^{th} century, physics seemed, for the most part, understood and complete. There were a few open questions but it appeared that is was the *mathematical* difficulty of solving equations that hindered answers. It came to the point that in 1874 Philipp von Jolly, a physics professor in Munich even disparaged Max Planck from studying physics because there were merely details left to solve.

Back then, physics distinguished between two categories of objects, matter (particles) and radiation (fields):

Matter	**Radiation**
• is made of particles that can be characterized at all points in time by their location $\vec{r}$ and their momentum $\vec{p}$	• is not made of particles but demonstrates wave-like behavior
• obeys the laws of classical mechanics (Newton, Lagrange, Hamilton)	• obeys Maxwell's equations

The atomic structure could not be proven directly but was confirmed indirectly in many ways. Because typical bodies are obviously made up of a lot of particles (1 mol consists of $N_L = 6.022 \times 10^{23}$ particles) one could not solve the classical equations of motion directly. Besides, the individual motion of each particle was not interesting, because it couldn't be singularly observed, instead it was more important to understand the macroscopic, averaged characteristics of matter. Hence, *thermodynamics* and *statistical physics* were developed. However, these fields did not yet challenge the concepts of classical physics.

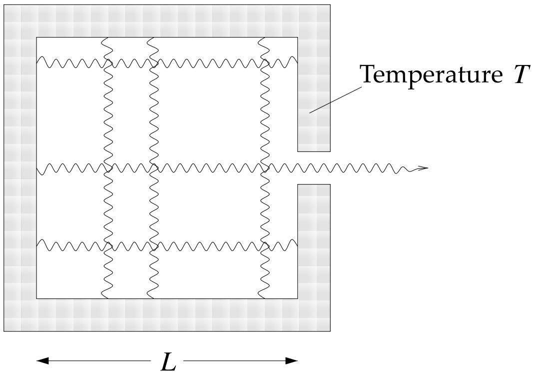
It was well known that particle systems could show wave-like phenomena, water waves and sound waves are prime examples for this. Therefore, it was natural to guess that there should be a material carrier also of electromagnetic waves, which was called *ether* back then. If this idea had been successful then radiation should have led back to matter in the table above and everything could have come to a large unification within classical physics. That idea was debunked by Michelson and Morley in 1887, who showed that the speed of light was independent of its direction relative to the earth's, and thus the lab's, motion. This could not be explained if earth were to move relative to an ether. After this discovery it seemed clear that matter and electromagnetic waves should be two separate categories. However, after some delay these experiments led to the first groundbreaking change at the dawn of the 20^{th} century: the special theory of relativity. For us it is important to understand that the special theory of relativity as well as the general theory of relativity did not yet go beyond the classical picture of exactly localized particles and wave-like radiation.

2.2 Black Body Radiation

A *black body* is defined as a body that absorbs the entire incident radiation. It is actually a misnomer because any body in a thermodynamic equilibrium with an electromagnetic field absorbs the exact amount it emits. A black body at high temperatures is not black, it glows.

The sun is within good approximation a black body, even though it is not in thermal equilibrium with its surroundings. But the plasma that we do see, the so-called photosphere of the sun, is approximately in equilibrium with deeper layers which provide heat to the photosphere. The emission spectrum is thus approximately that of a black body.

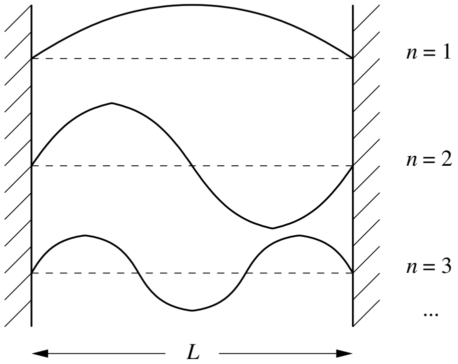
One experimental realization of a black body is a cavity with a small opening. Almost all of the light hitting the opening from the outside is absorbed, and nearly all of the light radiating from the small opening is emitted by the inner walls.



Within a classical picture, one can calculate the spectral energy density $w(\nu)$ of the outgoing light. Take a cubic, grounded metal box: the electromagnetic field inside can be decomposed into standing waves if we neglect the opening. Only standing waves whose parallel electric-field components $\vec{E}_{\parallel}$ and orthogonal induction-field components $B_{\perp}$ disappear at the surface are possible.

In one dimension we would have the condition:

$$n\,\frac{\lambda}{2} = L, \quad n = 1, 2, 3, \ldots \tag{2.1}$$



Therefore, with a wave number of $k = 2\pi/\lambda$,

$$k = n\,\frac{\pi}{L}, \quad n = 1, 2, 3, \ldots \tag{2.2}$$

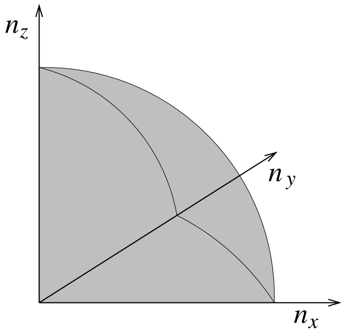
Hence, in a cubic cavity we would have the wave vector

$$\vec{k} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} \frac{\pi}{L}, \quad n = 1, 2, 3, \ldots \tag{2.3}$$

The frequency of the wave is

$$\nu = \frac{c}{\lambda} = \frac{c|\vec{k}|}{2\pi} = \frac{c}{2L}\sqrt{n_x^2+n_y^2+n_z^2} = \frac{c}{2L}\,|\vec{n}|. \tag{2.4}$$

How many field modes with frequencies smaller or equal to ν exist?



For $\sqrt{n_x^2+n_y^2+n_z^2} \gg 1$ or $L \gg \lambda$ we can neglect the discreteness of our points (n_x, n_y, n_z). Then there are

$$N(\nu) = \underbrace{2}_{\text{Polarizations}} \times \underbrace{\frac{1}{8}}_{\substack{\text{only one octant}\\(n_x,n_y,n_z>0)}} \times \frac{4\pi}{3}\Big(\underbrace{\frac{2L}{c}\,\nu}_{=|\vec{n}|}\Big)^3 = \frac{8\pi}{3}\,\frac{L^3}{c^3}\,\nu^3 \tag{2.5}$$

field modes with frequencies smaller or equal to ν. The number of modes with frequencies in the interval $[\nu, \nu+d\nu]$ is $dN = (dN/d\nu)\,d\nu$ with

$$\frac{dN}{d\nu} = 8\pi\,\frac{L^3}{c^3}\,\nu^2. \tag{2.6}$$

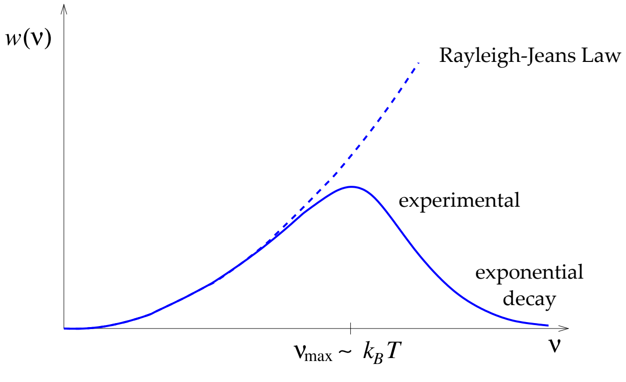
According to the equipartition theorem (see lecture on thermodynamics and statistical physics), every mode in equilibrium has the mean energy k_BT, to which the $\vec{E}$ and $\vec{B}$ field each contribute $k_BT/2$. The spectral energy density per volume is then

$$w(\nu) = \frac{1}{L^3}\,\frac{dN}{d\nu}\,k_BT = 8\pi\,\frac{\nu^2}{c^3}\,k_BT, \tag{2.7}$$

also called the *Rayleigh–Jeans law*. Within the framework of classical physics, it is exact. However, we can immediately see that it cannot be correct because the entire energy density is

$$w = \int_0^\infty d\nu\,w(\nu) = 8\pi\,\frac{k_BT}{c^3}\int_0^\infty d\nu\,\nu^2, \tag{2.8}$$

which diverges at higher frequencies (*ultraviolet catastrophe*). Also compared to experiment, classical theory fails for this part:



In 1900, Planck assumed that the standing electromagnetic waves were resonant with certain (not further specified) oscillators within the walls of the cavity and that these could only exist in states of certain *discrete* energies $E_n = n\varepsilon$, $n = 0, 1, 2, \ldots$ It was quite a remarkable assumption because the development of quantum theory would only commence a few years later—Niels Bohr published his atomic model in 1913. In statistical physics it is shown that the probability for energies E_n to occur is proportional to $e^{-E_n/k_BT} = e^{-n\varepsilon/k_BT}$ (also called the Boltzmann factor). This leads to the mean energy of such an oscillator,

$$\bar{E} = \frac{0\, e^{-0/k_BT} + \varepsilon\, e^{-\varepsilon/k_BT} + 2\varepsilon\, e^{-2\varepsilon/k_BT} + \ldots}{e^{-0/k_BT} + e^{-\varepsilon/k_BT} + e^{-2\varepsilon/k_BT} + \ldots} = \frac{\sum_{n=0}^{\infty} n\varepsilon\, e^{-n\varepsilon/k_BT}}{\sum_{n=0}^{\infty} e^{-n\varepsilon/k_BT}}, \tag{2.9}$$

which can also be written as

$$\bar{E} = \frac{\sum_{n=0}^{\infty} n\varepsilon\, e^{-n\beta\varepsilon}}{\sum_{n=0}^{\infty} e^{-n\beta\varepsilon}}, \tag{2.10}$$

with $\beta = 1/k_BT$. The infinite series can be resummed,

$$\begin{aligned} \bar{E} &= -\frac{d}{d\beta} \ln \sum_{n=0}^{\infty} e^{-n\beta\varepsilon} = -\frac{d}{d\beta} \ln \frac{1}{1 - e^{-\beta\varepsilon}} \qquad \Big| \text{(geometric series)} \\ &= \frac{d}{d\beta} \ln(1 - e^{-\beta\varepsilon}) = \frac{\varepsilon e^{-\beta\varepsilon}}{1 - e^{-\beta\varepsilon}} \\ &= \frac{\varepsilon}{e^{\beta\varepsilon} - 1} = \frac{\varepsilon}{e^{\varepsilon/k_BT} - 1}. \end{aligned} \tag{2.11}$$

Experiments show that ε must be proportional to the frequency ν. The constant of proportionality is now called *Planck's constant* h. So we can thus write ε as

$$\varepsilon = h\nu. \tag{2.12}$$

Planck replaced the mean energy k_BT in the Rayleigh–Jeans law by $\bar{E}$, leading to *Planck's law* (or Planck's radiation law)

$$w(\nu) = 8\pi \frac{\nu^3}{c^3} \frac{h}{e^{h\nu/k_BT} - 1}. \tag{2.13}$$

For $h\nu \ll k_BT$, it transitions into the Rayleigh–Jeans law formula (because $e^{h\nu/k_BT} - 1 \cong h\nu/k_BT$) and predicts an exponential decay for $h\nu \gg k_BT$. Note that Planck assumed discrete energy levels for the oscillators in the walls of the cavity. Nowadays, we know that the formula for radiation actually comes from the discrete occupation of electromagnetic-field modes, i.e., from the particle characteristics of light. The wall itself only functions as a heat bath with the temperature T.

2.3 The Photoelectric Effect

In 1887, Heinrich Hertz observed that electrons leave metal surfaces under irradiation with (UV) light. Thee details of this so-called *photoelectric effect* or for short *photoeffect* were experimentally researched by Wilhelm Hallwachs (professor at the TH Dresden from 1893–1922, rector from 1921–1922). Hertz and Hallwachs made the following observations:

1. The photoeffect only occurs when the frequency ν of the light surpasses a material dependent threshold frequency ν_g.
2. The kinetic energy of the emitted *photoelectrons* is proportional to $\nu - \nu_g$,

   $$\frac{1}{2} mv^2 \propto \nu - \nu_g \text{ for } \nu > \nu_g. \tag{2.14}$$

3. It is independent of the light intensity, defined as the transported energy per time per area orthogonal to the direction of propagation,

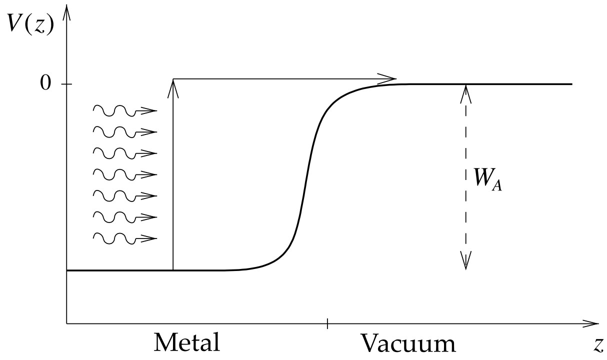
   $$I = \frac{1}{2} c\,\varepsilon_0\, E_0^2. \tag{2.15}$$

   Here, E_0 is the amplitude of the electric field component of the light wave.

4. The number of photoelectrons emitted per time, and thus the current, is proportional to the intensity I.
5. Emission follows immediately without any measurable delay, even at extremely low intensities.

Electromagnetic waves carry energy, described by the intensity I, as we know from classical electrodynamics. It seems plausible that this energy is transferred to the electrons by absorption from a metal surface, which causes their emission. So what does classical physics predict for these photoelectrons? Let's take two extreme cases:

1. If the energy transferred to an electron per period τ of the light wave is much smaller than the binding energy within the metal (also called the *work function* W_A), then the electron should be emitted after multiple periods, gathering more energy every period until it leaves the surface of our metal.

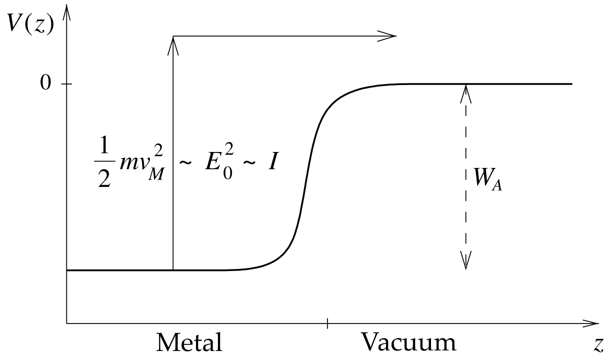


   The kinetic energy of the electrons should then be small since they barely have enough energy to leave the surface. Most importantly, the kinetic energy should be independent of the light frequency. This is in contradiction to the experiment. There should also be a delay between the absorption phase and emission phase, which also contradicts the experiment.

2. If the time necessary to transfer energy until emission is small relative to the period τ, then the force affecting the electrons should be proportional to the electric amplitude E_0 (the force is $\vec{F} = -e\vec{E}$). If the electrons were in a state of rest before absorption, then they should be accelerated to a velocity v_M (M for "metal") proportional to E_0. The energy relative to the ground state within the metal should then be

$$\frac{1}{2} m v_M^2 \propto E_0^2 \propto I \tag{2.16}$$

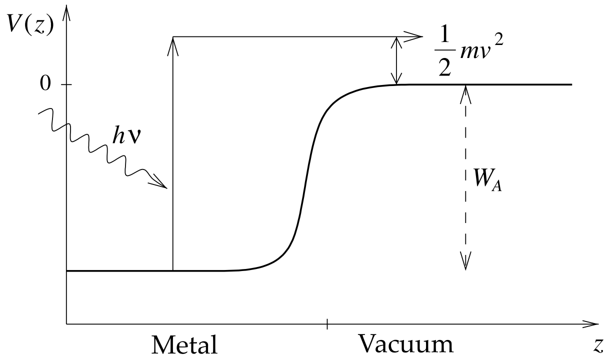
   in other words, also proportional to the light intensity I, which again contradicts the experiment. (After leaving the surface, the kinetic energy should be $(1/2)mv^2 = \alpha I - W_A$ with a constant α.)



The photoeffect was only correctly explained in 1903 by Albert Einstein, for which he was awarded the Nobel prize in physics in 1922 (for the year 1921). Einstein took a crutical further step that Planck had avoided and postulated that the electromagnetic radiation itself (and not just some oscillators in solid bodies as proposed by Planck) had to consist of bundles of energy with $E = h\nu$. Their number per time and per cross section area then had to be proportional to their intensity. According to Einstein, energy could only be transferred between the light and the metal surface via absorption or emission in complete packages (*light-quantum hypothesis*). These packages (light quanta) behaved as particles. However, in contrast to other particles such as electrons, their number does not have to be conserved. Today, these particles are called *photons*.

If an electron with a binding energy (work function) of W_A absorbs a photon with energy $h\nu$ the residual kinetic energy is

$$\frac{1}{2}\,mv^2 = h\nu - W_A. \tag{2.17}$$



For $(1/2)mv^2 > 0$, it must hold that $h\nu > W_A$, which implies

$$\nu_g = \frac{W_A}{h} \tag{2.18}$$

for the threshold frequency, which coincides quite nicely with the experiment. The photo-electron current is furthermore proportional to the number of absorbed photons per time and thus scales with the intensity. All theoretical results are well confirmed by experiments.

Because photons have the velocity c they cannot have a mass according to the special theory of relativity (STR). As for their momentum p, the theory predicts

$$E^2 = p^2c^2 + m^2c^4 = p^2c^2 \quad \Rightarrow \quad p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}. \tag{2.19}$$

With the angular frequency $\omega = 2\pi\nu$ and the wave number $k = 2\pi/\lambda$, one often writes

$$E = \frac{h\omega}{2\pi} \quad \text{und} \quad p = \frac{hk}{2\pi}. \tag{2.20}$$

And if one defines the abbreviation

$$\hbar := \frac{h}{2\pi} = 1{,}055 \times 10^{-34}\,\text{Js}, \tag{2.21}$$

the two relations become

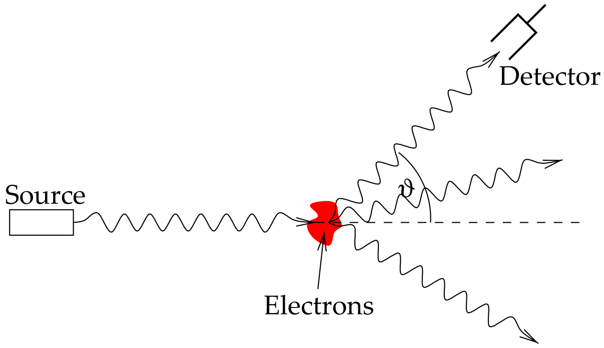
$$E = \hbar\omega \quad \text{and} \quad p = \hbar k. \tag{2.22}$$

2.4 The Compton Effect

The Compton effect is the elastic scattering of light off free or at least weakly bound electrons. Because it was only discovered in 1924 it couldn't contribute to the downfall of classical physics. However, it did confirm Einstein's particle model of electromagnetic radiation. Arthur H. Compton experimentally observed that upon scattering, the wavelength of light shifted by

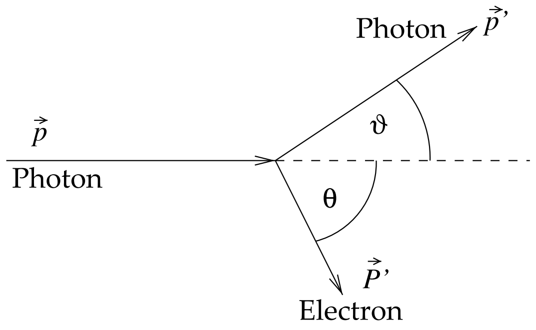
$$\Delta\lambda = 4\pi\,\frac{\hbar}{mc}\,\sin^2\frac{\vartheta}{2} \tag{2.23}$$

where ϑ is the deflection (or scattering) angle of the light.



It is important that $\Delta\lambda$ is independent of the wavelength λ for the incident beam, which implies it is independent of the frequency as well. Classically, one would expect a continuous energy transfer between light and the initially stationary electrons. This would mean a distribution of $\Delta\lambda$ for fixed scattering angle and not one single, sharp value of $\Delta\lambda$, as observed by Compton.

Within the particle model of photons, the experiment can be viewed as an elastic collision between a photon and an initially stationary electron. Momentum and energy are conserved for elastic collisions,



$$\vec{p} + 0 = \vec{p}\,' + \vec{P}', \tag{2.24}$$

$$pc + mc^2 = p'c + \underbrace{\sqrt{(P')^2c^2 + m^2c^4}}_{\text{(relativistic energy-momentum relation)}}. \tag{2.25}$$

The conservation of momentum gives

$$\begin{aligned} \vec{P}' = \vec{p} - \vec{p}\,' \quad \Rightarrow \quad (\vec{P}')^2 &= p^2 + (p')^2 - 2\vec{p}\cdot\vec{p}\,' \\ &= p^2 + (p')^2 - 2pp'\cos\vartheta. \end{aligned} \tag{2.26}$$

Insertion into the conservation of energy then yields

$$pc + mc^2 = p'c + \sqrt{p^2 + (p')^2 - 2pp'\cos\vartheta + m^2c^2}\; c \tag{2.27}$$

$$\Rightarrow \quad p - p' + mc = \sqrt{p^2 + (p')^2 - 2pp'\cos\vartheta + m^2c^2} \tag{2.28}$$

$$\Rightarrow \quad \cancel{p^2} + \cancel{(p')^2} + \cancel{m^2c^2} - 2pp' + 2pmc - 2p'mc = \cancel{p^2} + \cancel{(p')^2} - 2pp'\cos\vartheta + \cancel{m^2c^2}. \tag{2.29}$$

So with $p = h/\lambda$, we have

$$-\frac{2h^2}{\lambda\lambda'} + \frac{2mhc}{\lambda} - \frac{2mhc}{\lambda'} = -\frac{2h^2}{\lambda\lambda'}\cos\vartheta \tag{2.30}$$

$$\Rightarrow \quad -1 + \frac{mc}{h}\lambda' - \frac{mc}{h}\lambda = -\cos\vartheta \tag{2.31}$$

$$\Rightarrow \quad \Delta\lambda = \lambda' - \lambda = \frac{h}{mc}(1 - \cos\vartheta) = 2\frac{h}{mc}\sin^2\frac{\vartheta}{2} = 4\pi\frac{\hbar}{mc}\sin^2\frac{\vartheta}{2}. \tag{2.32}$$

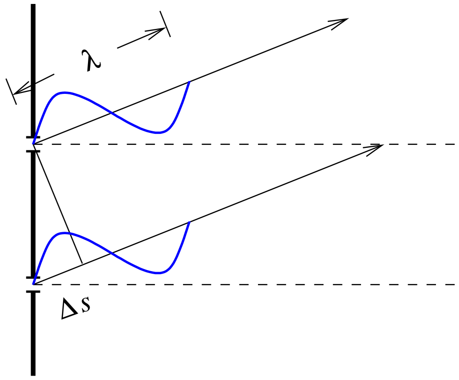
This is the observed shift. $\lambda_C := h/mc = 2.426 \times 10^{-12}\,\mathrm{m}$ is called the *Compton wavelength* of an electron and $\bar{\lambda}_C := \hbar/mc$ is the *reduced Compton wavelength*.

2.5 Wave or Particle?

The experiments discussed so far can only be explained naturally with a particle model of light. The particle picture of matter seemed obvious already. It had a long history with important experiments such as those performed by Thompson, Townsend, and Wilson in 1896, which addressed the elemental character of the electron as well as its charge and mass. In this section, we will discuss a few experiments that show that a simple particle model is insufficient.

2.5.1 Wave-Particle Duality of light

Certain aspects, such as the well-known diffraction and interference of light, can be successfully explained within the wave model of light. For example, in the double-slit experiment, approximately equidistant maxima and minima are observed if the path difference Δs between the two slits and a point on a far-away screen is equal to $\Delta s = n\lambda$ (or $\Delta = (n + 1/2)\,\lambda$) with $n = 0, \pm 1, \pm 2, \ldots$



Even if the intensity is reduced to a *single* photon at a time hitting the double slit the same interference pattern appears. Now, in a classical particle model, one would expect that because the particle can only go through one slit the other slit should be irrelevant. The image on the screen should simply be a superposition of the patterns produced by the two slits. But this is obviously not the case. One is forced to assume that a single photon interacts with *both* slits, even if there is a macroscopic distance between the two. This contradicts the classical prediction of particle trajectories.

We conclude that light behaves like a wave in some and like a collection of particles in other situations. This was called *wave-particle duality* but not understood at the time.

2.5.2 Wave-Particle Duality of Matter

In 1927/28, several independent experiments showed diffraction and interference of matter waves. More specifically, the reflection and transmission of electrons by crystals were investigated. The electron wavelength in these experiments was found to be

$$\lambda = \frac{h}{p} = \frac{h}{mv}, \tag{2.33}$$

the *de Broglie wavelength.* Note that it is always larger than but will converge towards the Compton wavelength in the relativistic limit,

$$\lambda = \frac{h}{mv} \overset{v<c}{>} \frac{h}{mc} = \lambda_C. \tag{2.34}$$

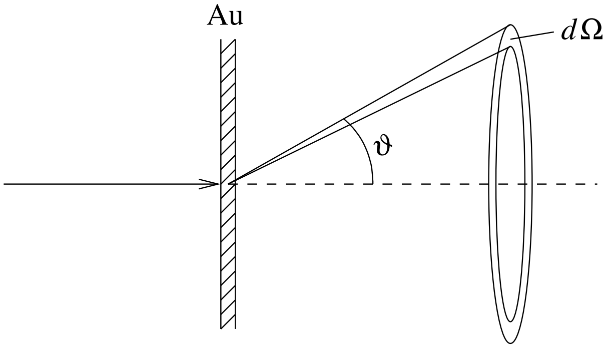
In this case, the theoretical prediction (de Broglie 1923) preceded the experimental confirmation.

So the experiments show that electrons also behave like particles in some and like waves in other situations. Later on, experiments demonstrated similar results for other particles such as neutrons. Nowadays, neutron diffraction is a standard method for the determination of crystal structures as well as of magnetic order.

2.6 The Failure of Rutherford's Model

Ernest Rutherford researched scattering of α particles (from radioactive decay) by thin gold foils from 1906 to 1913. The fraction of scattering into an solid-angle element $d\Omega$ with the diffraction angle ϑ relative to the incident direction can easily be evaluated using classical physics. One must simply assume nearly all of the atomic mass to be concentrated in a very small nucleus carrying charge $+Ze$. Then one can derive Rutherford's law of scattering

$$dN \propto N \, \frac{Z^2}{\sin^4 \frac{\vartheta}{2}} \, d\Omega. \tag{2.35}$$



At first sight, the agreement between the classical formula and the experimental results seems to confirm classical physics, specifically, classical mechanics and electrodynamics. Actually, a quantum mechanical approach will result in the same formula. This will be discussed in Quantum Theory 2. However, the experiments undermine classical physics in a different way, namely in the understanding of Rutherford's atomic model.

2.6.1 Stability of Atoms

According to Rutherford, atoms are comprised of electrons (light particles with a charge of $-e$) and a nucleus (heavy particles with a charge of $+Ze$). The electrons orbit the nucleus according to the laws of classical physics. Since the Coulomb potential scales anti-proportionally to r, $V(\vec{r}) \propto -1/r$, and attracts the electron this is a Kepler-type situation: The orbits should be elliptical. But according to classical electrodynamics, accelerated charges must also radiate electromagnetic waves. Electrons on elliptical orbits are of course accelerated, their acceleration is

$$\ddot{\vec{r}} = \frac{\vec{F}}{m} = -\frac{e\vec{E}}{m} = -\frac{1}{m}\,\frac{Ze^2}{4\pi\varepsilon_0 r^3}\,\vec{r} \neq 0. \tag{2.36}$$

The radiation reduces the electron's energy. Since its kinetic energy cannot be negative, the potential energy decreases so that the electron falls into the nucleus' Coulomb potential. Within the classical framework, calculating the electron orbit shows that it falls into the nucleus within a timescale of around 10^{-11} s. (What happens when the electron collides with the nucleus classical physics cannot predict. Assuming the charge is distributed continuously within the nucleus' volume, the electron's energy at least doesn't diverge to $-\infty$.)

The upshot is that classical physics predicts that a Rutherford atom collapses within 10^{-11} s. Atoms would then transition into small, electrically neutral particles (nuclei with captured electrons within them) by radiating electromagnetic waves, which obviously disagrees with all observations. In the classical picture there is nothing to stabilize the distance of about 10^{-10} m between electron and nucleus in solids and liquids. This implies that there would be no earth.

2.6.2 Atomic Spectra

If the Rutherford atom could avoid collapse by some unknown mechanism, then it should at least be able to absorb and emit any amount of energy, up to a certain maximum. Within the classical Kepler problem there should be no reason as to why some values of binding energy are preferred over others.

In 1885, Balmer observed that pure hydrogen gas upon stimulation by electron collisions only emits light of discrete wavelengths. Later, further infrared and ultraviolet wavelengths of emission were found. *All* of these wavelengths followed an empirical formula called the *Rydberg formula*,

$$\frac{\nu}{c} = \frac{1}{\lambda} = R\left(\frac{1}{n^2} - \frac{1}{m^2}\right), \tag{2.37}$$

with natural numbers $m > n$ and a factor of R with dimension of inverse length. It was also found that all of these predicted lines were indeed present in the observed spectra. For any n, an entire *series* was found with spectral lines for $m = n+1, n+2, n+3, \ldots$:

$n = 1$	Lyman series
$n = 2$	Balmer series
$n = 3$	Paschen series
$n = 4$	Bracket series
$n = 5$	Pfund series
...	...

The behavior is the same except for a different factor for other hydrogen-like ions with one valence electron each, such as He^{+}, Li^{2+}, etc. For more complicated ions and atoms, there are also series, albeit not as simple as those expressed by the Rydberg formula. But one principle is the same, the *Rydberg–Ritz combination principle*: The sum and difference of two frequencies in the spectrum are often also frequencies in the spectrum.

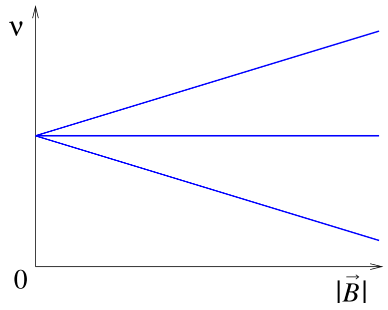
In the following chapter 3, we will return to the interpretation of spectra within the "old quantum theory" (Bohr, Sommerfeld). What can be said in advance is that, according to Einstein's light-quantum hypothesis, it makes sense that the spectrum should be interpreted as the emission of individual photons of energy $h\nu$. The natural interpretation of the relation

$$h\nu = \underbrace{hcR}_{=:\,E_R}\left(\frac{1}{n^2} - \frac{1}{m^2}\right) = -\frac{E_R}{m^2} + \frac{E_R}{n^2}, \tag{2.38}$$

together with the conservation of energy, is then that the emission is linked to a transition between two states of a hydrogen atom with energies of $-E_R/n^2$ and $-E_R/m^2$. E_R is called the *Rydberg energy* or, as a unit of energy, one *Rydberg*, where $1\,\mathrm{R} = 13.6\,\mathrm{eV}$. This argument leads to the conclusion that hydrogen atoms can only have discrete energies of $-E_R/n^2$, $n = 1, 2, \ldots$.

2.6.3 The Zeeman Effect

In an external magnetic field $\vec{B}$, the lines of the atomic spectra split into several lines. The splitting is proportional to $|\vec{B}|$ and was first observed by Zeeman in 1896.



The linear correspondence hints at the energy of a magnetic dipole $E_Z = -\vec{\mu} \cdot \vec{B}$ in a $\vec{B}$ field. Within Rutherford's atomic model, it is understandable, in principle, that the electron orbits around the nucleus should be connected to a magnetic dipole moment. An electron on a Kepler-type orbit creates a circular current. According to the laws of electrodynamics, this generates a magnetic moment perpendicular to the orbital plane.

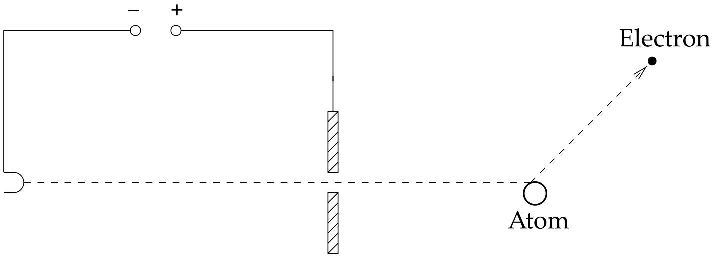
In the photon picture, it is natural to interpret the frequency shift as

$$\Delta\nu = \frac{\Delta E_Z}{h} = -\frac{\Delta\mu_z B}{h}, \tag{2.39}$$

with $\vec{B} = B\hat{e}_z$. The discreteness of the lines then shows that the z component of the dipole moment can only be changed by discrete values, coinciding with the later Stern–Gerlach experiment, see section 2.7.

2.6.4 The Franck–Hertz Experiment

The notion of discrete energy states of atoms was confirmed by Franck and Hertz in 1914. They researched the scattering of mono-energetic electrons off atoms.



Measurement of the kinetic energy of the electrons after scattering showed that the energy difference

$$\Delta E = E_{\text{before}} - E_{\text{after}} \tag{2.40}$$

only assumed discrete values $0 < \Delta E_1 < \Delta E_2 < \dots$ Especially if $E_{\text{before}} < \Delta E_1$, then the energy difference would be $\Delta E = 0$ since the smallest non-zero energy difference would have to be $\Delta E = \Delta E_1$ due to the impossibility of $E_{\text{after}} = E_{\text{before}} - \Delta E_1 < 0$. In these cases, the scattering thus had to be elastic.

Apparently, an atom only absorbs discrete energies $\Delta E_1, \dots$ It makes sense to view these energies as differences between two discrete energy states of the atom. The same energies were observed as in the optical spectrum. E.g., for a hydrogen atom, this means

$$\Delta E_1 = -\frac{E_R}{2^2} + \frac{E_R}{1^2} = \frac{3}{4}\,E_R, \tag{2.41}$$

$$\Delta E_2 = -\frac{E_R}{3^2} + \frac{E_R}{1^2} = \frac{8}{9}\,E_R \qquad \text{and so on.} \tag{2.42}$$

2.7 The Stern–Gerlach Experiment

Permanent magnets have been known to mankind since classical antiquity. Oersted's experiments in 1819 proved that electric currents generate magnetic fields. So when it came to permanent magnets, the easiest explanation was to assume the existence of electric currents within the material. Rutherford's atomic model provided a seemingly obvious explanation for currents within each singular atom due to the orbiting charged electrons. These currents should generate a magnetic dipole moment. Otto Stern and Walther Gerlach wanted to investigate the magnetic dipole moments of single atoms in 1921/22 and for this purpose sent a beam of neutral silver atoms through an inhomogeneous magnetic field. The energy of a magnetic dipole moment $\vec{\mu}$ within a $\vec{B}$ field is

$$E_{\text{Zeeman}} = -\vec{\mu}\cdot\vec{B}. \tag{2.43}$$

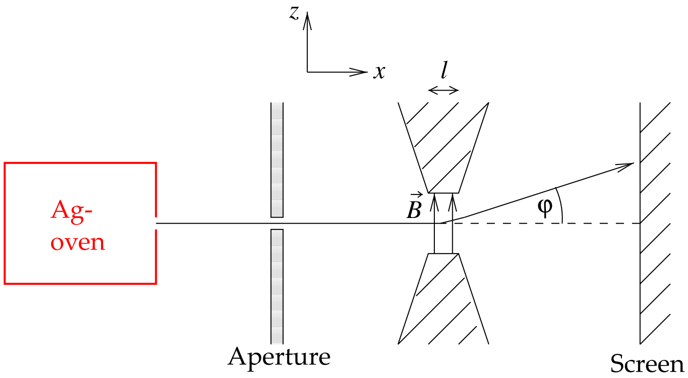
So if we choose the z-axis along $\vec{B}$, then

$$E_{\text{Zeeman}} = -\vec{\mu}\cdot\hat{z}\,B = -\mu_z B \tag{2.44}$$

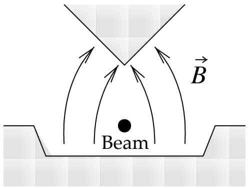
(we indicate unit vectors by a hat above the letter). The force acting on the dipole due to this potential energy is

$$\vec{F} = -\vec{\nabla} E_{\text{Zeeman}} = \mu_z \vec{\nabla} B, \tag{2.45}$$

which is only non-zero for an inhomogeneous magnetic field.



Cross section:



We will calculate the deflection of the atoms based on the assumption that the time spent in the $\vec{B}$ field is short. Then the atoms receive an impulse of force

$$\Delta\vec{p} = \vec{F}\Delta t \quad \text{with} \quad \Delta t = \frac{l}{v}. \tag{2.46}$$

In our case, v is the velocity of the incident atom. The velocity component in the z direction before the field region is zero and behind it is

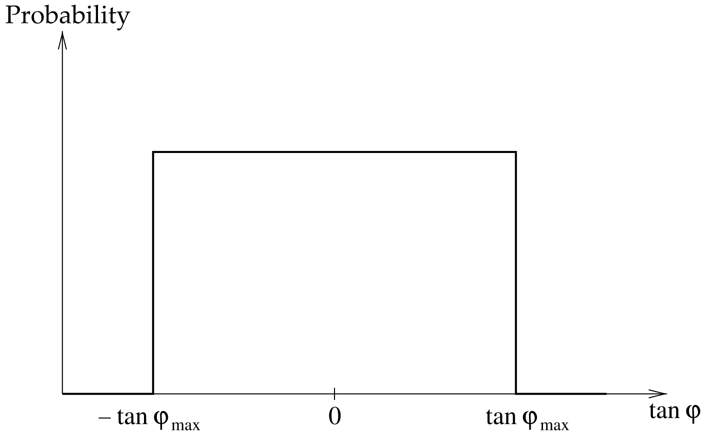
$$v_z = \frac{p_z}{m} = \frac{F_z}{m}\Delta t = \frac{\mu_z \frac{\partial B}{\partial z}}{m}\,\frac{l}{v}. \tag{2.47}$$

The diffraction angle φ fulfills

$$\tan\varphi = \frac{v_z}{v_x} = \frac{v_z}{v} = \frac{l}{mv^2}\,\frac{\partial B}{\partial z}\,\mu_z. \tag{2.48}$$

We assume that the silver atoms have a dipole moment of $\mu > 0$. When entering the $\vec{B}$ field, the direction of the dipole moment should be randomly distributed (isotropic) so that the z component μ_z should have a continuous distribution between $-\mu$ and μ. The expected distribution on the screen should accordingly be wide between the angles $-\varphi_{\max}$ and $\varphi_{\max}$ with

$$\tan\varphi_{\max} = \frac{l}{mv^2}\,\frac{\partial B}{\partial z}\,\mu. \tag{2.49}$$

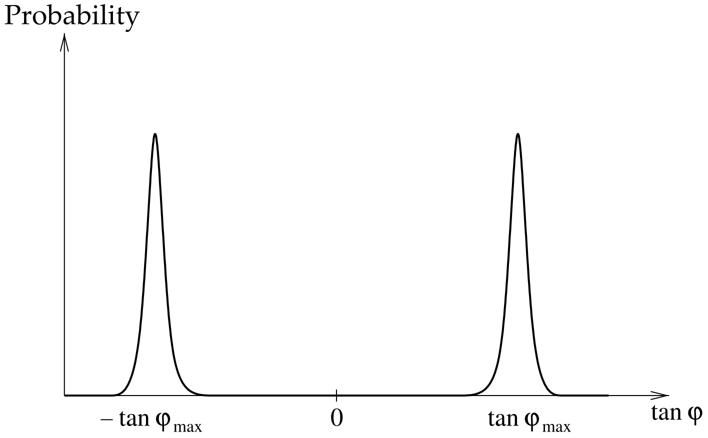


In fact, $\tan\varphi$ should be *equally distributed* because $\tan\varphi \propto \mu_z = \vec{\mu}\cdot\hat{z} = \mu\cos\theta$ (θ is the polar angle of the dipole moment in a polar coordinate system) and the distribution function of $\cos\theta$ is

$$\begin{aligned} P(\cos\theta) &= \frac{1}{4\pi}\int_0^{2\pi} d\phi' \int_0^{\pi} d\theta'\ \sin\theta'\,\delta(\cos\theta - \cos\theta') \\ &= \frac{1}{2}\int_0^{\pi} d\theta'\ \sin\theta'\,\delta(\cos\theta - \cos\theta') \qquad \Big|\ \text{substitute } u' = \cos\theta' \\ &= \frac{1}{2}\int_{-1}^{1} du'\,\delta(\cos\theta - u') = \frac{1}{2} = \text{const} \end{aligned} \tag{2.50}$$

for $\cos\theta \in [-1, 1]$.

But that isn't what comes out of the experiment. Instead, one finds only two sharp maxima, which both have the same height.



The broadening is based on the experimental resolution and is not fundamental in nature. Within a good approximation, the two observable diffraction angles are

$$\mu_z = \pm\mu_B := \pm\frac{e\hbar}{2m_e}, \tag{2.51}$$

with the electron mass m_e (*not* the mass of a silver atom). Apparently, the z component of the dipole moment can only assume two discrete values, $\pm\mu_B$, with μ_B being the Bohr magneton. This cannot be understood within classical physics. We will explain the discreteness observed in the experiments by Stern and Gerlach as well as Zeeman later on, based on the quantum mechanical characteristics of angular momentum.

For Ag^+ ions, no splitting is visible, so they don't have a magnetic dipole moment. Due to this, one can attribute the observed moment to the valence electron present in the outer shell of a neutral silver atom but missing in an ionized one.

Chapter 3

The Old Quantum Theory

The so-called “old quantum theory” was developed by Niels Bohr and others around 1913 to describe the apparent discrete energy values of atoms and to allow the calculation of atomic spectra. It is based on the assumption that classical physics can describe the dynamics of particle systems in principle but that only certain orbits are allowed. These allowed orbits are given by *ad hoc quantization rules.* We will see that Bohr’s quantum theory only describes closed orbits.

3.1 Bohr’s Postulates

The assumption that only certain orbits are possible contradicts classical physics, which predicts that a charged particle would have to radiate electromagnetic waves and thus continuously loose energy. This radiation must be avoided by some mechanism which Bohr never described. On the other hand, certain discrete frequencies of radiation were still observed. Bohr expressed his assumptions within two postulates:

1. Periodic orbits are only possible with specific discrete energies E_1, E_2, ... They do not radiate.
2. An electron can transition between two allowed periodic orbits with the energy E_n and E_m under emission or absorption of electromagnetic radiation with a frequency ν of

$$\nu = \frac{E_n - E_m}{h}. \tag{3.1}$$

This does not yet imply that a system with periodic orbits, such as an atom, has a ground state. (This cannot be concluded even if it is known that the energy has a lower bound. As a counterexample, take $E_n = E_1/n, n = 1, 2, \ldots$.) But physically relevant cases all have a ground state. When an atom is in its ground state it can’t radiate any energy so that it stays in this state until excited by external effects. We can easily interpret the frequency $\nu = (E_n - E_m)/h$ within the particle model of light. Using the Planck relation $\epsilon = h\nu$ reveals that a photon is absorbed or emitted under the transition. Note that a photon can also be “virtual” (in modern language), meaning that it can also be directly absorbed by another particle, such as by a free electron in the Franck–Hertz experiment.

3.2 Quantization Rules by Bohr and Sommerfeld

Bohr’s postulates don’t mention anything about *which* of the classical periodic orbits are actually allowed. We will discuss this just like Bohr, first for the hydrogen atom and then more generally.

3.2.1 The Hydrogen Atom

In classical mechanics, the attractive Coulomb interaction between a proton and electron results in a Kepler-type problem. For all negative energies $E < 0$, the orbits are elliptical and in the limiting case circular, each with the orbital frequency

$$\nu_{\text{orbit}} = \frac{1}{T_{\text{orbit}}} = \frac{1}{\pi}\,\frac{4\pi\varepsilon_0}{e^2}\sqrt{\frac{2|E|^3}{m}}, \tag{3.2}$$

where $m \approx m_e$ represents the reduced mass, and with the longer semiaxis

$$a = \frac{1}{2}\,\frac{e^2}{4\pi\varepsilon_0}\,\frac{1}{|E|}. \tag{3.3}$$

As discussed in chapter 1, classical physics describes macroscopic processes extremely well. Hence, quantum theory should coincide with classical physics in the limit large length scales, i.e., in this case for large a. With this argument, we are applying the important *correspondence principle* to the hydrogen atom. We will get back to this. According to Eqs. (3.2) and (3.3), large a implies small values of $|E|$ and thus low frequencies ν_{orbit}.

Classically speaking, an atom with a given energy E should radiate electromagnetic waves with a frequency of ν_{orbit} as well as higher harmonics $2\nu_{\text{orbit}}$, $3\nu_{\text{orbit}}$, ... (Higher harmonics are only radiated because an electron on an elliptical orbit isn't an ideal oscillating dipole.) Now, Bohr's postulates state that an atom radiates waves with frequencies of the form $(E_n - E_m)/h$. According to the correspondence principle, we assert that for small energies $|E| = |E_n|$ the smallest possible frequency ν_{orbit} (the fundamental frequency) is equal to the smallest possible quantum mechanical frequency,

$$\nu_{\text{orbit}} = \nu = \frac{E_n - E_{n-1}}{h}. \tag{3.4}$$

Since E_n grows with n—this is how we enumerate the allowed energies—and E_n is negative smaller values of $|E_n|$ correspond to larger quantum numbers n. If $E_n - E_{n-1} \ll |E_n|$, which we should still prove, we can also express the frequency as

$$\nu_{\text{orbit}} = \frac{1}{h}\,\frac{dE_n}{dn}. \tag{3.5}$$

According to the correspondence principle, this relation should hold for large n. Bohr then stipulated that it should be correct for *all* n. The equation

$$\frac{1}{h}\,\frac{dE_n}{dn} = \nu_{\text{orbit}}(E_n) \tag{3.6}$$

is then a differential equation for E_n as a function of n. Using separation of variables solves it with

$$\frac{dE_n}{\nu_{\text{orbit}}(E_n)} = h\,dn \tag{3.7}$$

and

$$\int_{-\infty}^{E_n} \frac{dE'_n}{\nu_{\text{orbit}}(E'_n)} = (n + n_c)\,h, \tag{3.8}$$

where n_c is an integration constant. Specifically for the hydrogen atom, we end up with

$$(n + n_c)\,h = \int_{-\infty}^{E_n} \frac{dE'_n}{\frac{1}{\pi}\,\frac{4\pi\varepsilon_0}{e^2}\sqrt{\frac{2|E'_n|^3}{m}}} = \frac{e^2}{4\varepsilon_0}\sqrt{\frac{m}{2}} \int_{-\infty}^{E_n} \frac{dE'_n}{(-E'_n)^{3/2}}. \tag{3.9}$$

We can evaluate the integral for $E_n < 0$:

$$(n + n_c)\,h = \frac{e^2}{4\varepsilon_0}\sqrt{\frac{m}{2}}\,2\,\frac{1}{\sqrt{-E'_n}}\Bigg|_{-\infty}^{E_n} = \frac{e^2}{2\varepsilon_0}\sqrt{\frac{m}{2}}\,\frac{1}{\sqrt{-E_n}}, \tag{3.10}$$

which gives

$$E_n = -\frac{e^4 m}{8\varepsilon_0^2 h^2}\,\frac{1}{(n + n_c)^2}. \tag{3.11}$$

We cannot infer the constant n_c from the correspondence principle. Comparing it to the observed spectrum shows no shift so that $n_c = 0$ if $n = 1, 2, 3, \ldots$ Then

$$E_n = -\frac{E_R}{n^2}, \tag{3.12}$$

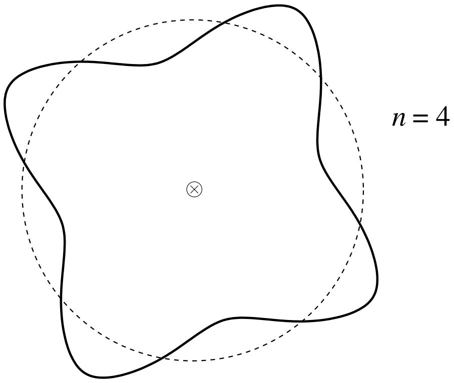
with

$$E_R = \frac{e^4 m}{8\varepsilon_0^2 h^2} = 2\pi^2 \left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 \frac{m}{h^2} = 13.6\,\text{eV}. \tag{3.13}$$

The differences between these quantized energies give the exact frequencies of the observed spectral lines via $\nu = (E_n - E_m)/h$. This was a huge success for Bohr's atomic model.

De Broglie gave a clear justification for the discrete energy values. The idea is to view the orbiting electrons as waves. A stable wave should only be possible if the orbit is a multiple of the electron's (de Broglie) wavelength since anything else would lead to destructive interference with an average amplitude of zero. We thus find that

$$2\pi r = n\lambda = n\,\frac{h}{p}, \quad n = 1, 2, 3, \ldots \tag{3.14}$$



Now we know that

$$V = -\frac{e^2}{4\pi\epsilon_0}\,\frac{1}{r}, \tag{3.15}$$

and, according to the virial theorem, the kinetic energy is

$$T = -\frac{V}{2} = +\frac{1}{2}\,\frac{e^2}{4\pi\epsilon_0}\,\frac{1}{r}. \tag{3.16}$$

The momentum follows as

$$p = \sqrt{2mT} = \sqrt{m\frac{e^2}{4\pi\epsilon_0}\frac{1}{r}}\,. \tag{3.17}$$

Equation (3.14) then gives

$$r = n\,\frac{\hbar}{p} = \frac{n\hbar}{\sqrt{m}}\,\sqrt{\frac{4\pi\epsilon_0}{e^2}}\,\sqrt{r} \tag{3.18}$$

$$\Rightarrow \qquad \sqrt{r} = \frac{\hbar}{\sqrt{m}}\,\sqrt{\frac{4\pi\epsilon_0}{e^2}}\,n \tag{3.19}$$

$$\Rightarrow \qquad r = \frac{\hbar^2}{m}\,\frac{4\pi\epsilon_0}{e^2}\,n^2, \tag{3.20}$$

and the energy is

$$E = T + V = -\frac{1}{2}\,\frac{e^2}{4\pi\epsilon_0}\,\frac{1}{r}$$

$$= -\frac{1}{2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \frac{m}{\hbar^2}\frac{1}{n^2}$$

$$= -2\pi^2\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \frac{m}{h^2}\frac{1}{n^2}, \quad n = 1, 2, \ldots \tag{3.21}$$

So we find exactly the same discrete energy values as predicted by Bohr's quantization rules. Equation (3.20) is also interesting: For large n the orbital radius grows to be much larger than the ground state radius, which is equal to $r = r_1 \approx 1\,\text{Å} = 10^{-10}\,\text{m}$. For large n we should thus reach the classical limit so that Eq. (3.20) should give us a good approximation for the size of atoms. Such comparatively huge atoms with large n are called *Rydberg atoms*. Since they can easily be deexcited or ionized by collisions their small life time demands extremely low densities such as in extremely rarefied gases, available only in lab conditions or in interstellar space.

3.2.2 Quantization of Phase Space

So far, we've only limited the allowed values of the larger semiaxis a and of the energy E to discrete values. However, the smaller semiaxis b must also be limited since otherwise we would still have infinitely many possible orbits for each E with different b. That cannot be correct. To solve this problem and to solve for different potentials, we need to find a more systematic theory for the quantization of phase-space trajectories. In this section, we will introduce such a theory.

Let us formulate Bohr's ideas as the quantization of periodic orbits $(\vec{q}, \vec{p})(t)$ in phase-space. This formulation is mainly due to Wilson and Sommerfeld. To start with, we will only work with periodic systems with *one* degree of freedom. In this case, the phase-space is two dimensional and can be parameterized by q, p. The quantization condition can then be written as

$$\int_{E_{\min}}^{E_n} \frac{dE'_n}{\nu_{\text{orbit}}(E'_n)} = (n + n_c)\,h, \tag{3.22}$$

where the lower limit $E_{\min}$ allows for the fact that the classical energy can have a lower bound for potentials other than the Coulomb potential. The left-hand side can be rewritten as an integral in phase space (utilizing the Hamilton function H as an expressive for the energy),

$$\int_{E_{\min}}^{E_n} \frac{dH}{\nu_{\text{orbit}}(H)} = \int_{E_{\min}}^{E_n} dH\, T_{\text{orbit}}(H) = \int_{E_{\min}}^{E_n} dH \int_0^{T_{\text{orbit}}(H)} dt = \iint_{H(q,p)\,\leq\, E_n} dq\,dp\, \underbrace{\begin{vmatrix} \frac{\partial H}{\partial q} & \frac{\partial H}{\partial p} \\ \frac{\partial t}{\partial q} & \frac{\partial t}{\partial p} \end{vmatrix}}_{\text{Jacobi determinant}} \quad . \tag{3.23}$$

According to the canonical equations of Hamilton mechanics, the determinant is

$$\begin{vmatrix} \frac{\partial H}{\partial q} & \frac{\partial H}{\partial p} \\ \frac{\partial t}{\partial q} & \frac{\partial t}{\partial p} \end{vmatrix} = \begin{vmatrix} -\frac{dp}{dt} & \frac{dq}{dt} \\ \frac{\partial t}{\partial q} & \frac{\partial t}{\partial p} \end{vmatrix} = \left| -\frac{dp}{dt}\frac{\partial t}{\partial p} - \frac{dq}{dt}\frac{\partial t}{\partial q} \right| = \left| \frac{\partial t}{\partial q}\frac{dq}{dt} + \frac{\partial t}{\partial p}\frac{dp}{dt} \right|.$$

Now we view $t = t(q, p)$ as a function of q and p. If q or p for their part depend on some variable x the chain rule gives

$$\frac{dt}{dx} = \frac{\partial t}{\partial q}\frac{dq}{dx} + \frac{\partial t}{\partial p}\frac{dp}{dx}. \tag{3.24}$$

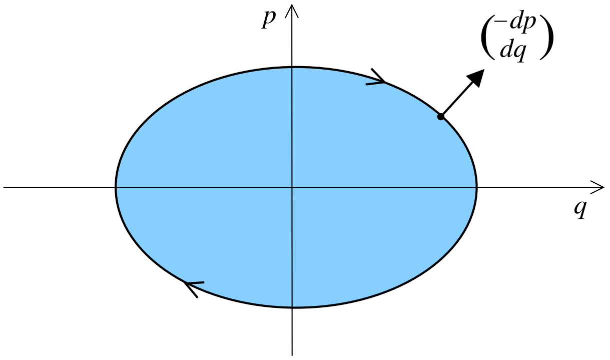
Now we set $x \equiv t$ and obtain, using the previous equation,

$$\begin{vmatrix} \frac{\partial H}{\partial q} & \frac{\partial H}{\partial p} \\ \frac{\partial t}{\partial q} & \frac{\partial t}{\partial p} \end{vmatrix} = \left| \frac{dt}{dt} \right| = 1. \tag{3.25}$$

From this we conclude that

$$\int_{E_{\min}}^{E_n} \frac{dH}{\nu_{\text{orbit}}(H)} = \iint_{H(q,p)\,\leq\, E_n} dq\,dp. \tag{3.26}$$

This is evidently the area in phase space enclosed by the orbit with the energy $E = E_n$.



This expression can also be rewritten as a line integral using Gauss' theorem (also known as the divergence theorem) in two dimensions,

$$\begin{aligned}\iint\limits_{H(p,q)\le E_n} dq\,dp &= \frac{1}{2}\iint\limits_{H\le E_n} dq\,dp\,\left(\frac{\partial}{\partial q},\frac{\partial}{\partial p}\right)\cdot\begin{pmatrix}q\\p\end{pmatrix}\\ &\overset{\text{Gauß}}{=} \frac{1}{2}\oint\limits_{H=E_n}(-dp,dq)\cdot\begin{pmatrix}q\\p\end{pmatrix} = \frac{1}{2}\oint\limits_{H=E_n}(p\,dq - q\,dp)\\ &= \frac{1}{2}\oint\limits_{H=E_n} p\,dq - \frac{1}{2}\Bigg[\underbrace{\oint\limits_{H=E_n} d(qp)}_{=\,qp\big|_{(q_0,p_0)}^{(q_0,p_0)}\,=\,0} - \oint\limits_{H=E_n} p\,dq\Bigg] = \oint\limits_{H=E_n} p\,dq. \end{aligned} \tag{3.27}$$

This provides a compact formulation of the quantization rule in phase space:

$$\oint\limits_{H(q,p)=E_n} p\,dq = (n+n_c)\,h, \quad n=1,2,\dots \tag{3.28}$$

The area enclosed by an orbit in phase space always has the dimension of action (energy times time), even for generalized coordinates. According to Bohr and Sommerfeld, it must be an integer multiple of Planck's constant up to a correction of $n_c h$, which defines the smallest possible orbit. It is easily recognizable that the rule is only applicable for periodic phase-space trajectories, and thus not for, e.g., scattering trajectories, since the integral path must be closed.

As an example, let's apply the method to a harmonic oscillator. In this case

$$H = \frac{p^2}{2m} + \frac{1}{2}\,m\omega_0^2 q^2. \tag{3.29}$$

For $H = E = \text{const}$, this is the equation for an ellipse with semiaxes

$$q_{\max} = \sqrt{\frac{2E}{m\omega_0^2}} \quad \text{and} \quad p_{\max} = \sqrt{2mE}. \tag{3.30}$$

The area of the ellipse is thus

$$\oint\limits_{H=E} p\,dq = \iint\limits_{H\le E} dq\,dp = \pi\,q_{\max}\,p_{\max} = \frac{2\pi E}{\omega_0} \overset{!}{=} (n+n_c)\,h. \tag{3.31}$$

We thus find

$$E_n = \hbar\omega_0\,(n+n_c). \tag{3.32}$$

The distances between the energy levels are correct, as we will see later. The constant can't be inferred from Bohr–Sommerfeld quantization. Schrödinger wave mechanics shows that we must choose $n_c = -1/2$ for $n = 1, 2, 3, \ldots$ or equivalently (and more commonly seen) $n_c = +1/2$ for $n = 0, 1, 2, \ldots$

The quantization rule can be applied to systems with multiple degrees of freedom if there are generalized coordinates $(\vec{q}, \vec{p})$ for which the equations of motions are decoupled for the various degrees of freedom. Should this be the case, the motion of (q_i, p_i) becomes independent of (q_j, p_j) for $i \neq j$. This is especially the case for so-called action-angle coordinates. Not all mechanical systems allow this. It is necessary and sufficient that the system is *integrable*, see the lecture on theoretical mechanics.

If the decoupling is possible, the phase-space trajectory must also be periodic for each pair (q_i, p_i) with potentially different periods and thus frequencies. Then, the quantization conditions are, according to Wilson and Sommerfeld,

$$\oint p_i \, dq_i = (n_i + n_i^c)\, h\,, \quad n_i = 1, 2, 3, \ldots \tag{3.33}$$

We apply this to the hydrogen atom, ignoring the constants n_i^c for now.

Since Kepler orbits are planar we are dealing with a two-dimensional problem. Polar coordinates r, φ and p_r, $p_\varphi = L$ are suitable generalized coordinates and conjugate momenta, respectively. Then the Hamilton is

$$H = \underbrace{\frac{1}{2m}\left(p_r^2 + \frac{L^2}{r^2}\right)}_{T} \underbrace{- \frac{e^2}{4\pi\varepsilon_0}\frac{1}{r}}_{V}. \tag{3.34}$$

For given energy $H = E = \text{const}$ and momentum $L = \text{const}$, conservation of energy stipulates that p_r only depends on r except for its sign but not on φ:

$$\frac{1}{2m}\left(p_r^2 + \frac{L^2}{r^2}\right) - \frac{e^2}{4\pi\varepsilon_0}\frac{1}{r} = E. \tag{3.35}$$

On the other hand, p_φ only depends on φ (in a most trivial manner: p_φ is a constant function of φ):

$$p_\varphi = L = \text{const}. \tag{3.36}$$

The two degrees of freedom are thus decoupled. The last relation leads to the quantization

$$\oint p_\varphi \, d\varphi = 2\pi L \overset{!}{=} n_\varphi h \tag{3.37}$$

$$\Rightarrow \quad L = n_\varphi \hbar, \quad n_\varphi = 1, 2, 3, \ldots \tag{3.38}$$

Equation (3.35) is more complicated and, after some calculation, leads to

$$\oint p_r \, dr = 2\pi \frac{e^2}{4\pi\varepsilon_0}\sqrt{-\frac{m}{2E}} - 2\pi L \overset{!}{=} n_r h, \quad n_r = 1, 2, \ldots \tag{3.39}$$

If we define the *azimuthal quantum number*, also known as the *orbital angular momentum quantum number*, as $l := n_\varphi = 1, 2, \ldots$ and the *principle quantum number* as $n := n_\varphi + n_r = 2, 3, \ldots$, we get

$$L = \hbar l \quad \text{und} \quad E = -\frac{E_R}{n^2}, \tag{3.40}$$

which is consistent with modern quantum mechanics. As we will see, the quantum numbers actually assume the values $n = 1, 2, 3, \ldots$ and $l = 0, 1, \ldots, n-1$. We already know that we could not find the constant n_i^c, so the beginning point of the integer numbers could not have been found. But the *upper* limit for the azimuthal quantum number is unexpected—again, the Bohr–Sommerfeld quantization rule cannot determine the possible quantum numbers and requires *ad hoc* assumptions.

For the sake of completeness, we provide the results for the semiaxes of the classical elliptical orbits:

$$a = \frac{1}{2}\frac{e^2}{4\pi\epsilon_0}\frac{1}{|E|} = \frac{\hbar^2}{m}\frac{4\pi\epsilon_0}{e^2}n^2, \tag{3.41}$$

see Eq. (3.20), and

$$b = \frac{l}{n}\,a. \tag{3.42}$$

Note that for $l = 0$, i.e., for angular momentum $L = 0$, the ellipse degenerates into a linear pendulum orbit through the nucleus, which can't be properly interpreted because of the singularity of $V(\vec{r})$ at $\vec{r} = 0$. But reality is not well described by Kepler orbits in the first place.

3.2.3 The Correspondence Principle

As the most general formulation of the old quantum theory we found that the action

$$S = \sum_i \oint p_i \, dq_i \tag{3.43}$$

of mechanical systems can only change by an integer multiple of a *quantum of action* h. At least, that it is what we've seen for periodic motions. Because of this we expect the behavior of a mechanical system to differ strongly from the classical predictions if the action S is on the order of magnitude of h. In contrast, the discreteness of action shouldn't have any effect if $S \gg h$. This is the main content of the *correspondence principle*: Quantum mechanics incorporates classical mechanics as a limiting case for $S \gg h$. We should thus be able to mathematically derive the laws of classical mechanics from the laws of quantum mechanics by taking the limit $h \to 0$ in a proper way. This correspondence should occur at the levels of observables. One example is Planck's radiation law as seen in section 2.2:

$$w(\nu) = 8\pi \frac{\nu^3}{c^3} \frac{h}{e^{h\nu/k_BT} - 1} \overset{h\to 0}{\cong} 8\pi \frac{\nu^3}{c^3} \frac{h}{h\nu/k_BT} = 8\pi \frac{\nu^2}{c^3} k_BT\,. \tag{3.44}$$

In the classical limit, we recover the Rayleigh–Jeans law, as we had already seen above.

3.3 Difficulties of the Old Quantum Theory

With the addition of an empirical value for the smallest allowed orbit, the old quantum theory allows a quantitative description of the spectra of hydrogen and of quite a number of other systems. The principle of quantization of phase space is so useful that it is still used today, namely to investigate quantum corrections for periodic motions that are much easier to calculate classically than quantum mechanically. Even so, the older theory still doesn't provide a satisfactory description and above all, it is incomplete:

- Additional empirical assumptions about allowed values of quantum numbers are required (constants n_i^c and sometimes excluded values, such as for the angular momentum).
- The description of classically unbound motions, such as scattering processes, isn't possible.
- The description of classically bound but *aperiodic* motions isn't possible either.
- The description of motion through *classically forbidden* regions (tunneling effect) isn't possible since there obviously isn't even a classical path in phase space to quantize.
- The spin and the anomalous Zeeman effect caused by the spin cannot be described.
- The theory also makes inconsistent predictions. For example, a hydrogen atom would have to fulfill the quantization condition $L/\hbar = 0, 1, 2, \ldots$ for *every* angular-momentum component so that

  $$\frac{\vec{L}\cdot\hat{e}}{\hbar} = 0, 1, 2, \ldots \tag{3.45}$$

  for all unit vectors $\hat{e}$. (This is based on the isotropy of space.) But this implies that $\vec{L} = 0$.

- The most important point of criticism is that the theory doesn't try to explain or justify the *ad hoc* quantization rules. From this perspective, the theory doesn't *explain* anything. On the contrary, the idea of classical orbits is in conflict with discontinuous transitions. The old quantum theory doesn't give any hints as to what "really happens" here.

Because of all these problems the old quantum theory alone isn't a satisfactory theory of quantum phenomena. Most of these problems are overcome by the more modern quantum theory, proposed and developed by Schrödinger, Heisenberg, etc.

Chapter 4

The Origins of Wave Mechanics

Modern quantum theory can be phrased in different ways. Wave mechanics is a very useful way of describing the motion of particles in space. Its foundation is Schrödinger's equation for the wave function. In this chapter, we will try to show that the structure of the Schrödinger equation is rather plausible. In the following chapter, we will then provide a more careful mathematical formulation of wave mechanics.

4.1 Matter Waves

In 1923, Louis de Broglie understood that light had both a particle-like (photo effect) and wave-like (diffraction) characteristics. He then suggested that the constituents of matter should also be subject to wave-particle duality. This seemingly daring idea was directly confirmed by diffraction experiments with electron beams. It allowed a very elegant interpretation of the quantization of closed orbits such as in the hydrogen atom, as we saw in section 3.2.1: On a closed orbit, we can only observe stationary waves if the circumference is an integer multiple of the wavelength. This is a general concept for waves or in general for fields and isn't confined to quantum theory—other examples from classical physics are the mechanical eigenmodes of a string or the electromagnetic eigenmodes of a cavity resonator.

The idea of matter waves begs the question of which equation of motion could describe these waves. The most important reference point for its construction is that again classical mechanics must be a limiting case of the theory (correspondence principle).

4.2 Inductive Reasoning for the Schrödinger Equation: Wave Packets

If we want to describe matter by waves we need to understand how to successfully return to the particle picture. One intuitive idea is to describe particles as *wave packets* with preferably small width in space and in momentum. However, we cannot localize the wave packet to an infinitely exact location and momentum: Let $\psi(\vec{k}, t)$ be a wave packet in real space, then its Fourier transform

$$\tilde{\psi}(\vec{k}, t) = \int d^3r \, e^{-i\vec{k}\cdot\vec{r}} \, \psi(\vec{r}, t) \tag{4.1}$$

is the corresponding distribution in momentum space ($\vec{k}$ space). If $\psi(\vec{r}, t)$ is narrow, then the corresponding $\tilde{\psi}(\vec{k}, t)$ must be broad, and if $\tilde{\psi}$ is narrow, then ψ must be broad. We sketch the underlying reasoning for the case of vanishing mean position $\bar{\vec{r}}$ and wave vector $\bar{\vec{k}}$. We will use the probability distributions $|\psi|^2$ and $|\tilde{\psi}|^2$ instead of ψ and $\tilde{\psi}$ in anticipation of Schrödinger's wave mechanics. Assuming this, the mean-square errors of position and momentum are

$$\Delta r^2 = \frac{\int d^3r \, r^2 \, |\psi(\vec{r}, t)|^2}{\int d^3r \, |\psi(\vec{r}, t)|^2}, \tag{4.2}$$

$$\Delta k^2 = \frac{\int \frac{d^3k}{(2\pi)^3}\, k^2\, |\tilde{\psi}(\vec{k},t)|^2}{\int \frac{d^3k}{(2\pi)^3}\, |\tilde{\psi}(\vec{k},t)|^2}. \tag{4.3}$$

We also assume that ψ is normalized,

$$\int d^3r\, |\psi(\vec{r},t)|^2 = 1. \tag{4.4}$$

It follows that

$$\begin{aligned}
\int \frac{d^3k}{(2\pi)^3}\, |\tilde{\psi}(\vec{k},t)|^2 &= \int \frac{d^3k}{(2\pi)^3} \int d^3r\, e^{i\vec{k}\cdot\vec{r}}\, \psi^*(\vec{r},t) \int d^3r'\, e^{-i\vec{k}\cdot\vec{r}\,'}\, \psi(\vec{r}\,',t) \\
&= \int d^3r\, d^3r'\, \delta(\vec{r}-\vec{r}\,')\, \psi^*(\vec{r},t)\, \psi(\vec{r}\,',t) \\
&= \int d^3r\, |\psi(\vec{r},t)|^2 = 1,
\end{aligned} \tag{4.5}$$

and we can conclude that

$$\Delta r^2 = \int d^3r\, r^2\, |\psi(\vec{r},t)|^2, \tag{4.6}$$

$$\Delta k^2 = \int \frac{d^3k}{(2\pi)^3}\, k^2\, |\tilde{\psi}(\vec{k},t)|^2. \tag{4.7}$$

Defining the auxiliary quantity

$$I(\lambda) := \int d^3r\, |\vec{r}\,\psi(\vec{r},t) + \lambda\,\vec{\nabla}\psi(\vec{r},t)|^2, \tag{4.8}$$

which is obviously positive for all λ, we can rewrite the integral as

$$\begin{aligned}
I(\lambda) &= \int d^3r \left[\psi^*\, r^2\, \psi + \psi^*\, \vec{r}\cdot\lambda\,\vec{\nabla}\psi + \lambda\,(\vec{\nabla}\psi^*)\cdot\vec{r}\,\psi + \lambda\,(\vec{\nabla}\psi^*)\cdot\lambda\,\vec{\nabla}\psi\right] \\
&\overset{\text{partiell}}{=} \int d^3r \left[\psi^*\, r^2\psi + \lambda\,\psi^*\,\vec{r}\cdot\vec{\nabla}\psi - \lambda\,\psi^*\vec{\nabla}\cdot(\vec{r}\,\psi) - \lambda^2\,\psi^*\,\nabla^2\psi\right] \\
&= \Delta r^2 - \lambda\int d^3r\,\psi^*\,(\underbrace{\vec{\nabla}\cdot\vec{r}}_{=3})\,\psi - \lambda^2\int d^3r\,\psi^*\,\nabla^2\psi \\
&= \Delta r^2 - 3\lambda - \lambda^2\int d^3r\,\psi^*\,\nabla^2\,\psi.
\end{aligned} \tag{4.9}$$

The last term can be written as

$$\begin{aligned}
\int d^3r\,\psi^*\,\nabla^2\psi &= \int d^3r \int \frac{d^3k\, d^3k'}{(2\pi)^6}\, e^{-i\vec{k}\cdot\vec{r}}\,\tilde{\psi}^*(\vec{k},t)\,\nabla^2 e^{i\vec{k}\,'\cdot\vec{r}}\,\tilde{\psi}(\vec{k}\,',t) \\
&= \int d^3r \int \frac{d^3k\, d^3k'}{(2\pi)^6}\, e^{-i\vec{k}\cdot\vec{r}}\,\tilde{\psi}^*(\vec{k},t)\,(-k'^2)\,e^{i\vec{k}\,'\cdot\vec{r}}\,\tilde{\psi}(\vec{k}\,',t) \\
&= \int \frac{d^3k\, d^3k'}{(2\pi)^6}\,(2\pi)^3\,\delta(\vec{k}'-\vec{k})\,\tilde{\psi}^*(\vec{k},t)\,(-k'^2)\,\tilde{\psi}(\vec{k}\,',t) \\
&= -\int \frac{d^3k}{(2\pi)^3}\, k^2\, |\tilde{\psi}(\vec{k},t)|^2 = -\Delta k^2.
\end{aligned} \tag{4.10}$$

It follows that

$$I(\lambda) = \Delta r^2 - 3\lambda + \lambda^2\,\Delta k^2 \geq 0 \tag{4.11}$$

$$\Rightarrow \quad \Delta r^2 \geq -\lambda^2\,\Delta k^2 + 3\lambda \quad \forall\lambda. \tag{4.12}$$

Because this is true for all λ it is true in particular for the value of λ that maximizes the right-hand side of the equation. This value of λ can be obtained by solving the equation

$$\frac{d}{d\lambda}\left(-\lambda^2\,\Delta k^2 + 3\lambda\right) = -2\lambda\,\Delta k^2 + 3 = 0 \tag{4.13}$$

$$\Rightarrow \quad \lambda = \frac{3}{2\,\Delta k^2}. \tag{4.14}$$

We conclude that

$$\Delta r^2 \geq -\frac{9}{4}\,\frac{1}{\Delta k^2} + \frac{9}{2}\,\frac{1}{\Delta k^2} = \frac{9}{4}\,\frac{1}{\Delta k^2} \tag{4.15}$$

$$\Rightarrow \quad \Delta r^2\,\Delta k^2 \geq \frac{9}{4} \tag{4.16}$$

$$\Rightarrow \quad \Delta r\,\Delta k \geq \frac{3}{2}. \tag{4.17}$$

One can reason that the factor of 3 in the numerator reflects the three spatial dimensions.

When we represent particles as wave packets it is natural to identify the position of the particle with the center of mass or the mean position of the wave packet,

$$\overline{\vec{r}} = \int d^3r\,\vec{r}\,|\psi(\vec{r},t)|^2. \tag{4.18}$$

Now we must find an equation for the dynamics of the wave $\psi(\vec{r},t)$ that fulfills the correspondence principle, i.e., the mean position $\overline{\vec{r}}$ should follow the predictions of classical mechanics in the limiting case.

The wavelength of a plane matter wave is $\lambda = h/p$ according to de Broglie's later confirmed prediction (see chapter 2.5). We can deduce that the wave number must then be

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{h}\,p \equiv \frac{p}{\hbar} \tag{4.19}$$

so that

$$p = \hbar k. \tag{4.20}$$

The most natural relation for the corresponding vectors would then be

$$\vec{p} = \hbar\vec{k}. \tag{4.21}$$

The mean momentum of a wave packet is then

$$\begin{aligned}
\overline{\vec{p}} \quad &= \quad \hbar\overline{\vec{k}} \;=\; \hbar\int \frac{d^3k}{(2\pi)^3}\,\vec{k}\,|\tilde{\psi}(\vec{k},t)|^2 \;=\; \hbar\int \frac{d^3k}{(2\pi)^3}\,\tilde{\psi}^*\,\vec{k}\,\tilde{\psi} \\
&= \quad \hbar\int \frac{d^3k}{(2\pi)^3}\int d^3r\,d^3r'\,e^{i\vec{k}\cdot\vec{r}}\,\psi^*(\vec{r},t)\,\vec{k}\,e^{-i\vec{k}\cdot\vec{r}\,'}\,\psi(\vec{r}\,',t) \\
&= \quad \hbar\int \frac{d^3k}{(2\pi)^3}\int d^3r\,d^3r'\,e^{i\vec{k}\cdot\vec{r}}\,\psi^*(\vec{r},t)\,\left(i\,\vec{\nabla}'e^{-i\vec{k}\cdot\vec{r}\,'}\right)\psi(\vec{r}\,',t) \qquad \Big|\;\vec{\nabla}'\text{ only affects }e^{-i\vec{k}\cdot\vec{r}\,'} \\
&\overset{\text{by parts}}{=} \hbar\int \frac{d^3k}{(2\pi)^3}\int d^3r\,d^3r'\,e^{i\vec{k}\cdot\vec{r}}\,\psi^*(\vec{r},t)\,e^{-i\vec{k}\cdot\vec{r}\,'}\,(-i\,\vec{\nabla}')\,\psi(\vec{r}\,',t) \\
&= \quad -i\hbar\int d^3r\,d^3r'\,\delta(\vec{r}-\vec{r}\,')\,\psi^*(\vec{r},t)\,\vec{\nabla}\psi(\vec{r}\,',t) \\
&= \quad -i\hbar\int d^3r\,\psi^*\,\vec{\nabla}\psi \\
&= \quad \int d^3r\,\psi^*\,\frac{\hbar}{i}\,\vec{\nabla}\psi.
\end{aligned} \tag{4.22}$$

This leads us to a heuristic rule: the momentum in position space is the “mean value” of the differential operator $\frac{\hbar}{i}\vec{\nabla}$. We also find analogous results for the mean of functions of momentum, e.g.,

$$\overline{p^2} = \hbar^2\overline{k^2} = \hbar^2 \int \frac{d^3k}{(2\pi)^3}\, \tilde{\psi}^*\, k^2\, \tilde{\psi} = -\hbar^2 \int d^3r\, \psi^*\, \nabla^2\psi = \int d^3r\, \psi^* \left(\frac{\hbar}{i}\vec{\nabla}\right)^2 \psi. \tag{4.23}$$

We thus assume that we should replace all momenta in integrals in position space by $\frac{\hbar}{i}\vec{\nabla}$. Then, the mean Hamilton function results in

$$\overline{H(\vec{r},\vec{p})} = \int d^3r\, \psi^* H\Big(\vec{r}, \frac{\hbar}{i}\vec{\nabla}\Big)\, \psi. \tag{4.24}$$

So now we need to construct an equation of motion for $\psi(\vec{r},t)$ such that the mean position $\bar{\vec{r}}$ and the mean momentum $\bar{\vec{p}}$ fulfill the classical equations of motion in the limiting case. First, let’s once again take the harmonic oscillator, now in three dimensions. Its classical Hamilton function is

$$H(\vec{r},\vec{p}) = \frac{p^2}{2m} + \frac{1}{2}\,\kappa r^2. \tag{4.25}$$

The canonical (Hamilton) equations of motion are

$$\dot{\vec{r}} = \frac{\partial H}{\partial \vec{p}} = \frac{\vec{p}}{m}, \tag{4.26}$$

$$\dot{\vec{p}} = -\frac{\partial H}{\partial \vec{r}} = -\kappa\vec{r}. \tag{4.27}$$

Thus we require that

$$\frac{d}{dt}\,\bar{\vec{r}} = \int d^3r\, \left(\dot{\psi}^*\, \vec{r}\psi + \psi^*\, \vec{r}\dot{\psi}\right) = \frac{1}{m}\int d^3r\, \psi^*\, \frac{\hbar}{i}\,\vec{\nabla}\psi = \frac{\bar{\vec{p}}}{m} \tag{4.28}$$

and

$$\frac{d}{dt}\,\bar{\vec{p}} = \int d^3r\, \left(\dot{\psi}^*\, \frac{\hbar}{i}\,\vec{\nabla}\psi + \psi^*\, \frac{\hbar}{i}\,\vec{\nabla}\dot{\psi}\right) = -\kappa \int d^3r\, \psi^*\, \vec{r}\psi = -\kappa\bar{\vec{r}}. \tag{4.29}$$

We rewrite the two equations as

$$\int d^3r\, \left(\dot{\psi}^*\, \vec{r}\psi + \psi^*\, \vec{r}\dot{\psi} - \psi^*\, \frac{1}{m}\,\frac{\hbar}{i}\,\vec{\nabla}\psi\right) = 0, \tag{4.30}$$

$$\int d^3r\, \left(\dot{\psi}^*\, \frac{\hbar}{i}\,\vec{\nabla}\psi + \psi^*\, \frac{\hbar}{i}\,\vec{\nabla}\dot{\psi} + \psi^*\, \kappa\vec{r}\psi\right) = 0. \tag{4.31}$$

As previously mentioned, we are searching for an equation of motion for ψ. It is reasonable to try to find an equation that is as simple as possible—we can still abandon the ansatz if it doesn’t work. Hence we will first attempt to find a linear equation that is of first order in time, i.e.,

$$\dot{\psi} = G\psi, \tag{4.32}$$

where G is an—as of yet an unknown—differential operator that can include functions of $\vec{r}$ and t as well as differentials of $\vec{r}$ (but not of t). Inserting G into the above equations gives

$$\int d^3r\, \left[(G\,\psi)^*\, \vec{r}\psi + \psi^*\, \vec{r}G\,\psi - \psi^*\, \frac{1}{m}\,\frac{\hbar}{i}\,\vec{\nabla}\psi\right] = 0, \tag{4.33}$$

$$\int d^3r\, \left[(G\,\psi)^*\, \frac{\hbar}{i}\,\vec{\nabla}\psi + \psi^*\, \frac{\hbar}{i}\,\vec{\nabla}\,G\,\psi + \psi^*\, \kappa\vec{r}\psi\right] = 0. \tag{4.34}$$

There are no systematic approaches for such a problem, so without educated guesses it is difficult to solve. We will make the ansatz that G is a sum of two terms, where one only depends on the nabla operator $\vec{\nabla}$ and the other only on $\vec{r}$,

$$G = G_1(\vec{\nabla}) + G_2(\vec{r}). \tag{4.35}$$

After partial integration for the G_1 parts we find that

$$\int d^3r\,\psi^*\left[G_1^*(-\vec{\nabla})\,\vec{r}+G_2^*(\vec{r})\,\vec{r}+\vec{r}\,G_1(\vec{\nabla})+\vec{r}\,G_2(\vec{r})-\frac{1}{m}\,\frac{\hbar}{i}\,\vec{\nabla}\right]\psi=0, \tag{4.36}$$

$$\int d^3r\,\psi^*\left[G_1^*(-\vec{\nabla})\,\frac{\hbar}{i}\,\vec{\nabla}+G_2^*(\vec{r})\,\frac{\hbar}{i}\,\vec{\nabla}+\frac{\hbar}{i}\,\vec{\nabla}G_1(\vec{\nabla})+\frac{\hbar}{i}\,\vec{\nabla}G_2(\vec{r})+\kappa\vec{r}\right]\psi=0. \tag{4.37}$$

This suggests that we set

$$G_1^*(-\vec{\nabla})=-G_1(\vec{\nabla}), \tag{4.38}$$

$$G_2^*(\vec{r})=-G_2(\vec{r}) \tag{4.39}$$

because we can then conclude that

$$\begin{aligned}&\int d^3r\,\psi^*\left[-G_1(\vec{\nabla})\,\vec{r}-\cancel{G_2(\vec{r})\,\vec{r}}+\vec{r}\,G_1(\vec{\nabla})+\cancel{\vec{r}\,G_2(\vec{r})}-\frac{1}{m}\,\frac{\hbar}{i}\,\vec{\nabla}\right]\psi\\ &\quad=\int d^3r\,\psi^*\left(-G_1\,\vec{r}+\vec{r}\,G_1-\frac{1}{m}\,\frac{\hbar}{i}\,\vec{\nabla}\right)\psi=0,\end{aligned} \tag{4.40}$$

$$\begin{aligned}&\int d^3r\,\psi^*\left[\cancel{-G_1(\vec{\nabla})\,\frac{\hbar}{i}\,\vec{\nabla}}-G_2(\vec{r})\,\frac{\hbar}{i}\,\vec{\nabla}+\cancel{\frac{\hbar}{i}\,\vec{\nabla}G_1(\vec{\nabla})}+\frac{\hbar}{i}\,\vec{\nabla}G_2(\vec{r})+\kappa\vec{r}\right]\psi\\ &\quad=\int d^3r\,\psi^*\left(-G_2\,\frac{\hbar}{i}\,\vec{\nabla}+\frac{\hbar}{i}\,\vec{\nabla}G_2+\kappa\vec{r}\right)\psi=0.\end{aligned} \tag{4.41}$$

The surviving terms of G_1 and G_2 don't vanish because we cannot naively change the order of $\vec{\nabla}$ and $\vec{r}$. We briefly show the next steps for the case of G_2:

1. A simple constant obviously doesn't solve Eq. (4.41):

$$G_2=ic,\quad c\in\mathbb{R}\quad\Rightarrow\quad\int d^3r\,\psi^*\,\kappa\vec{r}\,\psi=0,\quad\text{false!} \tag{4.42}$$

2. A linear solution doesn't work either:

$$G_2=i\vec{c}\cdot\vec{r},\quad\vec{c}\in\mathbb{R}^3 \tag{4.43}$$

$$\begin{aligned}\Rightarrow\quad&\int d^3r\,\psi^*\left[-i(\vec{c}\cdot\vec{r})\,\frac{\hbar}{i}\,\vec{\nabla}+\frac{\hbar}{i}\,\vec{\nabla}\,i(\vec{c}\cdot\vec{r})+\kappa\vec{r}\right]\psi\\ &=\int d^3r\,\psi^*\left(\cancel{-\hbar(\vec{c}\cdot\vec{r})\vec{\nabla}}+\hbar\vec{c}+\cancel{\hbar(\vec{c}\cdot\vec{r})\vec{\nabla}}+\kappa\vec{r}\right)\psi=0,\quad\text{false!}\end{aligned} \tag{4.44}$$

3. A quadratic function

$$G_2=icr^2,\quad c\in\mathbb{R} \tag{4.45}$$

solves the problem:

$$\begin{aligned}\int d^3r\,\psi^*\Big(-icr^2\,\frac{\hbar}{i}\,\vec{\nabla}+\frac{\hbar}{i}\,\vec{\nabla}icr^2+\kappa\vec{r}\Big)\psi&=\int d^3r\,\psi^*\Big(\cancel{-\hbar cr^2\vec{\nabla}}+\hbar c\,2\,\underbrace{r\hat{r}}_{=\vec{r}}\cancel{+\hbar cr^2\vec{\nabla}}+\kappa\vec{r}\Big)\psi\\ &=\int d^3r\,\psi^*\,(2\hbar c+\kappa)\,\vec{r}\,\psi=0.\end{aligned} \tag{4.46}$$

This requires $c=-\kappa/2\hbar$ so that

$$G_2=-\frac{i}{\hbar}\,\frac{\kappa}{2}\,r^2. \tag{4.47}$$

Analogously, we find

$$G_1 = i\hbar \, \frac{1}{2m} \nabla^2, \tag{4.48}$$

and thus the equation of motion is

$$\dot{\psi} = \frac{i\hbar}{2m} \, \nabla^2 \psi - \frac{i}{\hbar} \, \frac{\kappa}{2} \, r^2 \psi. \tag{4.49}$$

Slight rewriting gives

$$i\hbar \, \dot{\psi} = \left(-\frac{\hbar^2}{2m} \, \nabla^2 + \frac{\kappa}{2} \, r^2 \right) \psi. \tag{4.50}$$

This is the *Schrödinger equation* for the (three-dimensional) harmonic oscillator. This example also includes the case of a free particle for $\kappa = 0$. We have actually achieved more than expected: $\overline{\vec{r}}$ and $\overline{\vec{p}}$ *always* fulfill the classical equations. Our goal was that they fulfill the classical equations as a limiting case. Using the same approach for some other potential doesn't necessarily lead to an equation of motion for ψ that is consistent with the classical description of $\overline{\vec{r}}$ and $\overline{\vec{p}}$, as we will discuss later. It makes more sense to *postulate* the Schrödinger equation

$$i\hbar \, \dot{\psi} = \left(-\frac{\hbar^2}{2m} \, \nabla^2 + V(\vec{r}) \right) \psi \tag{4.51}$$

of the same form for an arbitrary potential $V(\vec{r})$. This is expected to be correct at least if the motion stays close to a normal minimum of $V(\vec{r})$ so that we can expand $V(\vec{r})$ up to second order around that minimum. But does it also hold in the classical limit? By that we mean the limit of a wave packet which is narrow both in position and in wave vector. We can view it as narrow in position if $V(\vec{r})$ (the only other position-dependent function in the problem) barely changes on a length scale given by the width Δr of the wave packet. But in that case we can locally expand $V(\vec{r})$ up to second order and Eq. (4.51) is thus justified. (Similarly, the kinetic energy should also be approximately quadratic in $\vec{p}$ for the classical limit case but we have assumed this anyway.)

For sure, we haven't actually derived the Schrödinger equation, we have at best shown that it is a plausible approach. Because quantum mechanics is the more general theory that includes classical mechanics as a limiting case it is impossible to deduce the equation from classical mechanics alone.

4.3 Inductive Reasoning for the Schrödinger equation: Hamilton–Jacobi Theory

In this section, we will present an alternative inductive reasoning for wave mechanics. To do so, we will need to recapitulate some concepts from within classical electrodynamics and mechanics.

4.3.1 The Eikonal Equation

The well-known wave equation is

$$\nabla^2 F - \frac{1}{c^2} \, \frac{\partial^2 F}{\partial t^2} = 0. \tag{4.52}$$

In electrodynamics, F could be a component of the $\vec{E}$ or $\vec{B}$ field. For a constant speed (of light) c, the equation is solved by plane waves,

$$F(\vec{r}, t) = F_0 \, e^{i(\vec{k}\cdot\vec{r} - \omega t)} = F_0 \, e^{i(\vec{k}\cdot\vec{r} - kct)}. \tag{4.53}$$

If the speed is not spatially constant the equation changes to

$$\nabla^2 F - \frac{n^2(\vec{r})}{c^2} \, \frac{\partial^2 F}{\partial t^2} = 0, \tag{4.54}$$

where $n(\vec{r})$ is the diffraction index. We are interested in the case of a gradually (spatially) changing $n(\vec{r})$, where "gradually" only has meaning in relation to the wavelength λ: $|\vec{\nabla} n|/n \ll 1/\lambda = k/2\pi$. We now make the ansatz

$$F(\vec{r}, t) = F_0 \, \exp\left(i\omega \left[\frac{L(\vec{r})}{c} - t \right] \right). \tag{4.55}$$

We call $L(\vec{r})$ the *optical path* or the *eikonal*. Insertion into the wave equation gives

$$\vec{\nabla} \cdot i\,\frac{\omega}{c}\,(\vec{\nabla} L) \exp\left(i\omega\left[\frac{L}{c} - t\right]\right) - \frac{n^2}{c^2}\,\frac{\partial}{\partial t}\,(-i\omega)\,\exp\left(i\omega\left[\frac{L}{c} - t\right]\right) = 0 \tag{4.56}$$

$$\Rightarrow \quad i\,\frac{\omega}{c}\,(\nabla^2 L)\,\cancel{\exp(\ldots)} - \frac{\omega^2}{c^2}\,(\vec{\nabla} L)^2\,\cancel{\exp(\ldots)} + \frac{n^2\omega^2}{c^2}\,\cancel{\exp(\ldots)} = 0 \tag{4.57}$$

$$\Rightarrow \quad (\vec{\nabla} L)^2 = n^2 + \frac{ic}{\omega}\,\nabla^2 L. \tag{4.58}$$

We can see that L must generally be complex, which is incompatible with the interpretation as an optical path. This is the result of assuming the amplitude to be constant.

We haven't used the premise of n only gradually changing spatially. Equation (4.58) can be solved for constant n by linear functions

$$L = L_0 + n\,\hat{e}\cdot\vec{r} \tag{4.59}$$

with arbitrary but fixed unit vectors $\hat{e}$. As a non-linear partial differential equation, Eq. (4.58) will generally have other solutions as well but the linear functions are the physically correct solutions for $n = \text{const}$ because they result in

$$F(\vec{r},t) = F_0\,\exp\left(i\omega\left[\frac{L_0 + n\,\hat{e}\cdot\vec{r}}{c} - t\right]\right) = F_0\,\underbrace{e^{i\omega L_0/c}}_{=\,\text{const}}\,e^{i(\vec{k}\cdot\vec{r}-\omega t)}, \tag{4.60}$$

with $\vec{k} = (\omega n/c)\,\hat{e}$. For such solutions, we find that

$$\nabla^2 L = \vec{\nabla}\cdot n\,\hat{e} = 0. \tag{4.61}$$

Hence, for constant n, the last term in Eq. (4.58) disappears. If n only gradually changes, then $\nabla^2 L$ should at least be small, and with it the last term as well. Then we can neglect this term compared to n^2 and arrive at the *eikonal equation*

$$(\vec{\nabla} L)^2 = n^2. \tag{4.62}$$

As mentioned, this equation is only applicable for a spatially gradually changing velocity. This is exactly the condition for the validity of *geometrical optics* or *ray optics*. (It is also applicable to cases with discontinuous jumps, which are generally solved by connecting solutions in areas of smooth changes to each other at the jumps, using appropriate continuity conditions.)

Equations for the form $L(\vec{r}) = \text{const}$ describe surfaces in space. On the other hand, the condition of a constant phase $\omega\,(L/c - t) = \text{const}$ describes wave fronts. Obviously, the wave fronts traverse the family of surfaces $L(\vec{r}) = \text{const}$ over time. The *rays* in geometrical optics are orthogonal at every point to the surfaces $L(\vec{r}) = \text{const}$ and thus also to the wave fronts.

4.3.2 Review of Hamilton–Jacobi Theory

The basic concept of Hamilton–Jacobi theory in classical mechanics is to find a transformation of the given canonical variables $(\vec{q},\vec{p})$ to new variables $(\vec{Q},\vec{P})$ which trivializes the transformed Hamilton function. To do so, one chooses a generating function of the type $F_2(\vec{q},\vec{P},t)$ (see lecture on theoretical mechanics), which will be denoted by S in our context, so that

$$p_i = \frac{\partial S}{\partial q_i}, \tag{4.63}$$

$$Q_i = \frac{\partial S}{\partial P_i}, \tag{4.64}$$

$$\mathcal{H} = H + \frac{\partial S}{\partial t}. \tag{4.65}$$

We should choose S in such a way that all new coordinates Q_i are cyclic, i.e., do not appear in the transformed Hamilton function $\mathcal{H}$. Then, all momenta P_i are conserved because of

$$\dot{P}_i = -\frac{\partial \mathcal{H}}{\partial Q_i} = 0. \tag{4.66}$$

Since the Q_i don't appear in $\mathcal{H}$ and the P_i generally appear in $\mathcal{H}$ but are conserved the only possible time dependence of $\mathcal{H}$ is explicit. Equation (4.65) shows that if we have any S that makes all Q_i cyclic, then we can add another (solely) t-dependent function to it in such a way that $\mathcal{H}$ doesn't depend on t anymore and even vanishes identically, $\mathcal{H} \equiv 0$.

The equation for the generating function with the desired properties, i.e., the *Hamilton–Jacobi equation*, is obtained from $\mathcal{H} \equiv 0$ as well as Eqs. (4.63) and (4.65),

$$H\left(q_1, \ldots, q_s, \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_s}, t\right) + \frac{\partial S}{\partial t} = 0, \tag{4.67}$$

where s is the number of degrees of freedom. It is a first-order partial differential equation for S as a function of $\vec{q}$ and t but is generally non-linear in the $\partial S/\partial q_i$. We will call the solutions for S *waves of action* for reasons we will explain later on. Since we would like to describe matter as a field (matter waves), the first important step is already solved with this approach: the Hamilton–Jacobi theory is a *field theory* for classical mechanics, namely for the field S. As discussed in the lecture on theoretical mechanics, there is a problem with the Hamilton–Jacobi theory that is often ignored: After the transformation we have s conserved momenta P_i. Since they are canonical momenta, they fulfill the relation

$$\{P_i, P_j\} = 0 \tag{4.68}$$

with the Poisson brackets

$$\{A, B\} := \sum_i \left(\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i}\right). \tag{4.69}$$

One says that the P_is are *in involution.* So we can see there are s conserved quantities in involution. But the number of conserved quantities in involution is independent of the choice of canonical variables and in particular remains unchanged by the transformation with S. Since not every mechanical system has s conserved quantities in involution it cannot always be possible to find a generating function S so that $\mathcal{H} \equiv 0$ holds true. Systems that *do* have s conserved quantities in involution are called *integrable.*

On the other hand, the Hamilton–Jacobi equation seems quite harmless, regardless of the question of integrability (which cannot easily be determined by looking at the equation). We cannot discern any reason as to why it shouldn't be solvable for at least almost all $(\vec{q}, \vec{p}, t)$. The resolution of this apparent paradox is that one typically cannot find solutions for non-integrable systems that are both unique and continuous in every point. We will now assume the solution to be unique and continuous in at least the region of interest.

We now consider the special case of Hamilton functions H without explicit time dependence. In this case, we can separate the time dependency from S using

$$S(\vec{q}, \vec{P}, t) = W(\vec{q}, \vec{P}) - Et. \tag{4.70}$$

Insertion into the Hamilton–Jacobi equation gives

$$H\left(q_1, \ldots, q_s, \frac{\partial W}{\partial q_1}, \ldots, \frac{\partial W}{\partial q_s}\right) = E. \tag{4.71}$$

Since the P_i are constant, equations with a structure such as $W = \text{const}$ define surfaces in configuration space, i.e., the space of $\vec{q}$. Surfaces with $S = \text{const}$, i.e., wave fronts of the waves of action S, push themselves past the stationary surfaces $W = \text{const}$ over time. The velocity of the wave fronts can be derived,

$$S = \text{const} \quad \Rightarrow \quad 0 = dS = \frac{\partial W}{\partial \vec{q}} \cdot d\vec{q} - E\, dt \tag{4.72}$$

so that we find

$$\frac{\partial W}{\partial \vec{q}} \cdot \vec{u} = E, \tag{4.73}$$

with $\vec{u}(\vec{r}) := d\vec{q}/dt$. Equation (4.73) doesn't uniquely determined $\vec{u}$—it is only a scalar equation but *vecu* has s components. We *define* $\vec{u}$ as orthogonal to the surfaces $W = \text{const}$ and thus to the wave fronts $S = \text{const}$ since any motion parallel to these surfaces doesn't describe a real change of S. The wave velocity $\vec{u}$ is then parallel to the gradient $\partial W/\partial\vec{q}$. Because of this parallelism we can conclude that

$$\left(\frac{\partial W}{\partial \vec{q}}\right)^2 u^2 = E^2 \tag{4.74}$$

$$\Rightarrow \quad \left(\frac{\partial W}{\partial \vec{q}}\right)^2 = \frac{E^2}{u^2}. \tag{4.75}$$

This is only a useful equation to determine W if we know $E^2/u^2(\vec{q})$ explicitly, which isn't always the case. Let us consider the following special case:

$$H = \frac{\vec{p}\cdot\vec{p}}{2m} + V(\vec{q}), \tag{4.76}$$

where h shouldn't be explicitly time dependent. This Hamilton function describes any number of particles of the same mass and non-relativistic kinetic energies with arbitrary one- and multi-particle potentials. It follows that for this Hamilton function

$$p^2 = 2m\,(H - V) = 2m\,(E - V). \tag{4.77}$$

Now the gradient $\partial W/\partial\vec{q}$ in Eq. (4.75) is exactly the vector of all momenta, $\vec{p}$. We can thus rewrite the equation as

$$\left(\frac{\partial W}{\partial \vec{q}}\right)^2 = 2m\,[E - V(\vec{q})]. \tag{4.78}$$

This equation together with Eq. (4.75) has the exact same structure as the *eikonal equation* (4.62).

The classical particle trajectories are the *rays* of the waves of action. These are the curves within the configuration space of $\vec{q}$ which are orthogonal to the surfaces $W = \text{const.}$

4.3.3 The Wave Equation of Quantum Mechanics

For electromagnetic waves, the eikonal equation was determined from the wave equation by taking the limit of geometrical optics. The central idea is to view the formally analogous eikonal equation (4.75) of classical mechanics,

$$\left(\frac{\partial W}{\partial \vec{q}}\right)^2 = \frac{E^2}{u^2}, \tag{4.79}$$

as a geometrical-optics limiting case of a more general type of mechanics, namely of quantum mechanics. As previously mentioned, the classical trajectories are seen as rays n this limiting case. Now we need to determine a wave equation that results in the eikonal equation in the limit of geometrical optics.

In analogy to the ansatz

$$F(\vec{r},t) = F_0 \,\exp\left(i\omega\left[\frac{L}{c} - t\right]\right) \tag{4.80}$$

for electromagnetic waves, we choos the *wave function* for matter waves as

$$\psi(\vec{q},t) = \psi_0\,\exp\left(i\underbrace{\omega}_{=E/\hbar}\left[\frac{W}{E} - t\right]\right) = \psi_0\,\exp\left(\frac{i}{\hbar}\,[W - Et]\right) = \psi_0\,e^{iS/\hbar}. \tag{4.81}$$

The constant $\hbar$ with the dimension of an action if required to make the phase dimensionless. But the value of $\hbar$ isn't given anywhere within the eikonal equation (or within the entire realm of classical physics). This is similar to how the electromagnetic eikonal equation $(\vec{\nabla}L)^2 = n^2$ doesn't contain the speed of light c, which is, however, an essential part of the the wave equation.

Which wave equation is satisfied by ψ? Note that the equations

$$\frac{\partial}{\partial\vec{q}}\psi = \frac{i}{\hbar}\,\frac{\partial W}{\partial \vec{q}}\,\psi, \tag{4.82}$$

$$\frac{\partial}{\partial t}\,\psi = -\frac{i}{\hbar}\,E\,\psi \tag{4.83}$$

hold. If we now apply an operator-valued function $G(\vec{q}, \partial/\partial\vec{q})$ of coordinates and coordinate derivatives to ψ, then every derivative $\partial/\partial q_i$ pulls a factor of $(i/\hbar)\,\partial W/\partial q_i$ in front of the exponential function. But the application of G does more since the derivatives act on everything that follows: Products with multiple factors of $\partial/\partial q_i$ generate higher derivatives of W. For example, we consider the function

$$H\left(\vec{q}, \frac{\hbar}{i}\,\frac{\partial}{\partial\vec{q}}\right) = -\frac{\hbar^2}{2m}\,\frac{\partial}{\partial\vec{q}}\cdot\frac{\partial}{\partial\vec{q}} + V(\vec{q}). \tag{4.84}$$

Applying H to ψ gives

$$H\left(\vec{q}, \frac{\hbar}{i}\,\frac{\partial}{\partial\vec{q}}\right)\psi = \frac{1}{2m}\,\frac{\partial W}{\partial\vec{q}}\cdot\frac{\partial W}{\partial\vec{q}}\,\psi - i\,\frac{\hbar}{2m}\,\underbrace{\left(\frac{\partial}{\partial\vec{q}}\cdot\frac{\partial}{\partial\vec{q}}W\right)}_{=\nabla_{\vec{q}}^2 W}\psi + V(\vec{q})\,\psi. \tag{4.85}$$

If $V(\vec{q})$ changes only gradually or jumps as a function of $\vec{q}$, then we can neglect the second derivatives of W with the same reasoning as in the derivation of the electromagnetic eikonal equation. So we see that the geometrical-optics limit yields

$$H\left(\vec{q}, \frac{\hbar}{i}\,\frac{\partial}{\partial\vec{q}}\right)\psi = \frac{1}{2m}\left(\frac{\partial W}{\partial\vec{q}}\right)^2\psi + V(\vec{q})\,\psi. \tag{4.86}$$

For the *same* Hamilton function H but with different arguments

$$H(\vec{q}, \vec{p}) = \frac{\vec{p}\cdot\vec{p}}{2m} + V(\vec{q}), \tag{4.87}$$

the Hamilton–Jacobi equation is

$$H\left(\vec{q}, \frac{\partial W}{\partial\vec{q}}\right) = \frac{1}{2m}\left(\frac{\partial W}{\partial\vec{q}}\right)^2 + V(\vec{q}) = E, \tag{4.88}$$

which in the limiting case of geometrical optics gives

$$i\hbar\,\frac{\partial\psi}{\partial t} \overset{\text{Gl. (4.83)}}{=} E\,\psi \overset{\text{Gl. (4.88)}}{=} H\left(\vec{q}, \frac{\partial W}{\partial\vec{q}}\right)\psi \overset{\text{Gl. (4.86)}}{=} H\left(\vec{q}, \frac{\hbar}{i}\,\frac{\partial}{\partial\vec{q}}\right)\psi. \tag{4.89}$$

With that, we have found the Schrödinger equation

$$i\hbar\,\frac{\partial\psi}{\partial t} = H\left(\vec{q}, \frac{\hbar}{i}\,\frac{\partial}{\partial\vec{q}}\right)\psi. \tag{4.90}$$

We evidently replaced all momenta p_i in the Hamilton function by coordinate derivatives, or, to be exact, by $(\hbar/i)\,\partial/\partial q_i$.

If we follow the same chain of logic backwards we find that the eikonal equation of classical mechanics, and with it essentially the Hamilton–Jacobi form of mechanics, is simply the geometrical-optics limit of the dynamics of matter waves as described by the Schrödinger equation. The inductive reasoning from a limiting case to a more general theory is, of course, not generally cogent and is closer to a postulate. Schrödinger further postulated that the equation is generally applicable, not only for the form of Hamilton functions we've considered thus far.

4.4 Further Remarks on the Schrödinger Equation

The Schrödinger equation for a massive particle with non-relativistic kinetic energy in the absence of a magnetic field is

$$i\hbar\,\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\,\nabla^2\psi + V(\vec{r})\,\psi. \tag{4.91}$$

We will discuss solutions of this equation in detail later. As mentioned, the right-hand side of the equation, except for the wave function ψ, arises from the classical Hamilton function by replacing the momentum $\vec{p}$ by $\frac{\hbar}{i}\vec{\nabla}$. Within a conservative system, as we assume ours to be, the Hamilton function is equal to the energy. This gives rise to the idea of understanding the left-hand side of the Schrödinger equation $i\hbar\,\frac{\partial}{\partial t}$ as a representation of energy. To replace

$$E \to i\hbar\,\frac{\partial}{\partial t}, \tag{4.92}$$

$$\vec{p} \to -i\hbar\,\vec{\nabla} \tag{4.93}$$

is plausible from a relativistic perspective since it can be summed up as the covariant rule

$$p^\mu \to i\hbar\,\partial^\mu. \tag{4.94}$$

However, this isn't a very strong argument since Schrödinger's theory isn't actually covariant. In any case, one must keep in mind that Eqs. (4.92) and (4.93) are just short-hand notations for the transition from classical to quantum mechanics. Quantum theory is not derived here and is only justified by the correspondence principle (classical mechanics is recovered a as a limiting case) and by comparison to experiments.

Let us briefly examine the case of vanishing potential,

$$i\hbar\,\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\,\nabla^2\psi. \tag{4.95}$$

It is solved by plane waves

$$\psi(\vec{r},t) = \psi_0\,e^{i(\vec{k}\cdot\vec{r}-\omega t)}. \tag{4.96}$$

Insertion yields

$$\hbar\omega\,\psi = \frac{\hbar^2 k^2}{2m}\,\psi, \tag{4.97}$$

and thus

$$\hbar\omega = \frac{\hbar^2 k^2}{2m}. \tag{4.98}$$

This gives the *dispersion relation*

$$\omega(\vec{k}) = \frac{\hbar k^2}{2m} \tag{4.99}$$

of free matter waves (compare the dispersion relation of free electromagnetic waves, $\omega(\vec{k}) = kc$). Assuming de Broglie's postulates of $E = \hbar\omega$ and $\vec{p} = \hbar\vec{k}$, the relation is consistent with the classical kinetic energy

$$E = \frac{p^2}{2m}. \tag{4.100}$$

Incidentally, Schrödinger first wrote down an equation with the structure

$$\left(\frac{\hbar^2}{c^2}\,\frac{\partial^2}{\partial t^2} - \hbar^2\,\nabla^2 + m^2c^4\right)\,\psi(\vec{r},t) = 0. \tag{4.101}$$

This equation doesn't provide the desired quadratic dispersion but rather

$$\omega^2(\vec{k}) = k^2c^2 + \frac{m^2c^4}{\hbar^2}. \tag{4.102}$$

Restructured toward energy and momentum, the equation reads

$$E^2 = p^2c^2 + m^2c^4, \tag{4.103}$$

which is known as the relativistic energy-momentum relation. Equation (4.101) is the so-called *Klein–Gordon equation*, which actually does help define relativistic quantum theory, see lecture on quantum theory 2. Schrödinger discarded this equation because it was of second order in regard to time.

Chapter 5

Schrödinger Wave Mechanics

5.1 Operators

Schrödinger's equation

$$i\hbar \frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{\hbar^2}{2m}\,\nabla^2\psi(\vec{r},t) + V(\vec{r})\,\psi(\vec{r},t) \tag{5.1}$$

contains terms that are generated by taking derivatives of the *wave function* $\psi(\vec{r},t)$ or by multiplying it with other functions. These terms can be unified by introducing the notion of *operators*. An operator $\hat{A}$ is a mapping from a certain space R, in our case a space of functions, into the *same* space:

$$\hat{A}: \begin{array}{ccc} R & \to & R \\ \psi & \mapsto & \hat{A}\psi. \end{array} \tag{5.2}$$

Note that the image of ψ is conventionally written as $\hat{A}\psi$ and not as $\hat{A}(\psi)$. In this lecture we will indicate operators using a hat (circumflex). Within quantum theory, we will predominantly deal with *linear* operators. They are defined by the condition that

$$\hat{A}(\lambda_1\,\psi_1 + \lambda_2\,\psi_2) = \lambda_1\,\hat{A}\psi_1 + \lambda_2\,\hat{A}\psi_2 \qquad \forall\,\psi_1,\psi_2 \in R \quad \text{and} \quad \lambda_1,\lambda_2 \in \mathbb{C}. \tag{5.3}$$

Two types of linear operators are especially important for wave mechanics:

1. Differential operators such as
$$\frac{\partial}{\partial x}:\ \psi(\vec{r},t) \mapsto \frac{\partial}{\partial x}\,\psi(\vec{r},t), \tag{5.4}$$
2. Multiplication operators such as
$$\hat{V}(\vec{r}):\ \psi(\vec{r},t) \mapsto V(\vec{r})\,\psi(\vec{r},t). \tag{5.5}$$

Operators can also behave as a vector; a vector operator is a vector the components of which are operators in the sense explained above. Important examples are

1. the Del operator, also called nabla,
$$\vec{\nabla} := \begin{pmatrix}\partial/\partial x\\ \partial/\partial y\\ \partial/\partial z\end{pmatrix}; \quad \vec{\nabla}:\ \psi(\vec{r},t) \mapsto \begin{pmatrix}\partial\psi/\partial x\\ \partial\psi/\partial y\\ \partial\psi/\partial z\end{pmatrix}, \tag{5.6}$$
2. multiplication with a vector, for example
$$\hat{\vec{r}}:\ \psi(\vec{r},t) \mapsto \vec{r}\,\psi(\vec{r},t). \tag{5.7}$$

It is easy to prove the following for linear operators $\hat{A}$, $\hat{B}$:

1. $c\hat{A}$ with a number $c \in \mathbb{C}$ is a linear operator,
2. $\hat{A} + \hat{B}$ is a linear operator, and
3. $\hat{A}\hat{B}$ is a linear operator, where the product is defined as applying the operators from right to left,
$$\hat{A}\hat{B}\,\psi(\vec{r},t) := \hat{A}[\hat{B}\psi(\vec{r},t)]. \tag{5.8}$$

It is crucial to understand that the product of operators is generally *non-commutative*. For example,

$$\hat{A} = \frac{\partial}{\partial x}, \quad \hat{B} = x, \quad \psi(x) = x^2. \tag{5.9}$$

This leads to

$$\hat{A}\hat{B}\,\psi = \frac{\partial}{\partial x}\,(x\,x^2) = \frac{\partial}{\partial x}\,x^3 = 3x^2, \tag{5.10}$$

whereas

$$\hat{B}\hat{A}\,\psi = x\,\frac{\partial}{\partial x}\,x^2 = x \times 2x = 2x^2. \tag{5.11}$$

One defines the *commutator* as

$$[\hat{A},\hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{5.12}$$

The commutator $[\hat{A},\hat{B}]$ is also an operator on the same space as $\hat{A}$ and $\hat{B}$. It represents a measure of non-commutativity. From the definition we can deduce a number of useful rules:

1. The commutator is antisymmetric under exchange of the arguments:
$$[\hat{B},\hat{A}] = \hat{B}\hat{A} - \hat{A}\hat{B} = -(\hat{A}\hat{B} - \hat{B}\hat{A}) = -[\hat{A},\hat{B}]. \tag{5.13}$$

2. The commutator is additive in both of its arguments:
$$[\hat{A}+\hat{B},\hat{C}] = (\hat{A}+\hat{B})\hat{C} - \hat{C}(\hat{A}+\hat{B}) = \hat{A}\hat{C} + \hat{B}\hat{C} - \hat{C}\hat{A} - \hat{C}\hat{B} = [\hat{A},\hat{C}] + [\hat{B},\hat{C}], \tag{5.14}$$
$$[\hat{A},\hat{B}+\hat{C}] = \hat{A}(\hat{B}+\hat{C}) - (\hat{B}+\hat{C})\hat{A} = \hat{A}\hat{B} + \hat{A}\hat{C} - \hat{B}\hat{A} - \hat{C}\hat{A} = [\hat{A},\hat{B}] + [\hat{A},\hat{C}]. \tag{5.15}$$

3. Commutators containing a product satisfy
$$\begin{aligned}[\hat{A}\hat{B},\hat{C}] &= \hat{A}\hat{B}\hat{C} - \hat{C}\hat{A}\hat{B}\\ &= \hat{A}\hat{B}\hat{C}\underbrace{- \hat{A}\hat{C}\hat{B} + \hat{A}\hat{C}\hat{B}}_{=\,0} - \hat{C}\hat{A}\hat{B}\\ &= \hat{A}(\hat{B}\hat{C} - \hat{C}\hat{B}) + (\hat{A}\hat{C} - \hat{C}\hat{A})\hat{B}\\ &= \hat{A}[\hat{B},\hat{C}] + [\hat{A},\hat{C}]\hat{B}.\end{aligned} \tag{5.16}$$

   The order of the factors is important since they don't generally commute. Using the commutator's antisymmetry leads to
$$[\hat{A},\hat{B}\hat{C}] = \hat{B}[\hat{A},\hat{C}] + [\hat{A},\hat{B}]\hat{C}. \tag{5.17}$$

   However, if one of the operators is the multiplication by a number $\lambda \in \mathbb{C}$ it follows that
$$[\lambda\hat{A},\hat{B}] = \lambda[\hat{A},\hat{B}] + \underbrace{[\lambda,\hat{B}]}_{=\,0}\hat{A} = \lambda[\hat{A},\hat{B}], \tag{5.18}$$
$$[\hat{A},\lambda\hat{B}] = \lambda[\hat{A},\hat{B}] + \underbrace{[\hat{A},\lambda]}_{=\,0}\hat{B} = \lambda[\hat{A},\hat{B}]. \tag{5.19}$$

   Together with the second property this implies that the commutator is linear in both of its arguments.

Example: For any differentiable function $\psi(x)$ it holds that

$$\left[\frac{\partial}{\partial x}, x\right]\,\psi(x) = \frac{\partial}{\partial x}\,x\,\psi(x) - x\,\frac{\partial}{\partial x}\,\psi(x) = \psi(x) + x\,\frac{\partial}{\partial x}\,\psi(x) - x\,\frac{\partial}{\partial x}\,\psi(x) = \psi(x). \tag{5.20}$$

Since this is true for any differentiable ψ we can write the result as the *operator identity*

$$\left[\frac{\partial}{\partial x}, x\right] = 1, \tag{5.21}$$

on the space of differentiable functions.

Heuristic reasoning in sections 4.2 and 4.3 led us to the *quantization rules*

$$\vec{r} \to \hat{\vec{r}} = \vec{r}, \tag{5.22}$$

$$\vec{p} \to \hat{\vec{p}} = \frac{\hbar}{i}\,\vec{\nabla}. \tag{5.23}$$

This leads to

$$[\hat{r}_i, \hat{p}_j] = \left[r_i, \frac{\hbar}{i}\,\frac{\partial}{\partial r_j}\right] = \frac{\hbar}{i}\,\left[r_i, \frac{\partial}{\partial r_j}\right] = -\frac{\hbar}{i}\,\underbrace{\left[\frac{\partial}{\partial r_j}, r_i\right]}_{\delta_{ij}} = i\hbar\,\delta_{ij}. \tag{5.24}$$

This *canonical commutation relation* is central for quantum theory.

Using the quantization rules also gives us the *Hamilton operator* or *Hamiltonian* $\hat{H}$, stemming from the Hamilton function,

$$H(\vec{r}, \vec{p}) \to \hat{H} = H\big(\hat{\vec{r}}, \hat{\vec{p}}\big). \tag{5.25}$$

The quantization rules do not always uniquely determine the the Hamilton operator since position and momentum do not commute. If the Hamilton function includes terms such as

$$H_1 = c\,\vec{r}\cdot\vec{p} = c\,\vec{p}\cdot\vec{r}, \tag{5.26}$$

with $c = \text{const}$, then it isn't obvious which of the following operators should be chosen:

$$\hat{H}_1 = \begin{cases} c\,\hat{\vec{r}}\cdot\hat{\vec{p}}, \\ c\,\hat{\vec{p}}\cdot\hat{\vec{r}}, \\ c\,\dfrac{\hat{\vec{r}}\cdot\hat{\vec{p}} + \hat{\vec{p}}\cdot\hat{\vec{r}}}{2}. \end{cases} \tag{5.27}$$

Or is the correct operator none of the three? Later on, we will see that the last choice, the symmetric form, is correct. This question is actually not a real problem since quantum mechanics is the more general theory so that we cannot expect an unambiguous deduction starting from classical mechanics.

In full generality, the classical Hamilton function depends on generalized coordinates q_j and their respective momenta p_j and possibly on time:

$$H = H(q_1, \ldots, p_1, \ldots, t). \tag{5.28}$$

According to section 4.3, we would like to generalize the quantization rules to

$$\begin{aligned} q_j &\overset{?}{\to} \hat{q}_j = q_j, \\ p_j &\overset{?}{\to} \hat{p}_j = \frac{\hbar}{i}\,\frac{\partial}{\partial q_j}. \end{aligned} \tag{5.29}$$

But these rules are contradictory. For example, we wouldn't find the same Hamilton operator if we applied the rules to cartesian or spherical coordinates describing the same system. Comparison to experiments and the consideration that coordinate unit vectors and with them their quantization rules should be independent of position leads to the conclusion that we must only use *cartesian* coordinates, as Dirac emphasized. Quantum mechanics can formulated for generalized coordinates but the quantization rules are always rooted in the fundamental canonical commutation relation for cartesian coordinates. Quantum mechanics in Schrödinger's formulation sets us back to the same level as Newton mechanics, which is also primarily expressed in cartesian coordinates.

5.2 The Schrödinger Equation as a Partial Differential Equation

The general Schrödinger equation for a single particle is

$$i\hbar \frac{\partial}{\partial t}\, \psi(\vec{r},t) = \hat{H}\, \psi(\vec{r},t), \tag{5.30}$$

where $\hat{H}$ is the Hamilton operator. $\hat{H}$ includes derivatives in terms of position, which means that it is a *partial* differential equation. Let us briefly discuss important characteristics of this equation as well as their consequences.

- The Schrödinger equation is *linear*. From this we can infer the *superposition principle*: If ψ_1 and ψ_2 are two solutions of the equation, then
  $$\psi = \lambda_1\, \psi_1 + \lambda_2\, \psi_2 \tag{5.31}$$
  is also a solution, for any constants $\lambda_1, \lambda_2 \in \mathbb{C}$. This is not only crucial for practical solutions for specific systems—many solution methods only work for linear equations—it is of tremendous fundamental importance. We will come back to this point when we discuss interpretations of quantum mechanics.

- The Schrödinger equation is of *first order* in time. Hence, for a given $\hat{H}$, specifying $\psi(\vec{r}, t_0)$ for all $\vec{r}$ and a certain time t_0 as the initial conditions uniquely determines the solution for all times $t > t_0$ (as well as $t < t_0$).

- If $\hat{H}$ includes second-order (or higher) spatial derivatives, such as for
  $$\hat{H} = \frac{\hat{p}^2}{2m} + V(\vec{r}), \tag{5.32}$$
  then the Schrödinger equation cannot be covariant, i.e., it contradicts special relativity, since to be covariant it would have to contain derivatives only in terms of the four-gradient
  $$(\partial_\mu) = \left(\frac{1}{c}\, \frac{\partial}{\partial t}, \vec{\nabla}\right). \tag{5.33}$$
  Since the equation of first order in time, it must also be of first order with in position. Dirac succeeded in constructing a covariant quantum mechanics, which will be discussed in the lecture quantum theory 2.

5.3 Probability Waves

We have yet to examine what the wave function $\psi(\vec{r}, t)$ actually means. If we assume a wave packet as an initial condition, then we typically find that its width Δx increases with time. The harmonic oscillator is an exception, though. For general potentials, the wave packet dissolves. Within the interpretation of classical particles in terms of localized wave packets from section 4.2, this leads to the expectation that the particle nature is gradually lost. However, we observe that particles do not dissolve over time—scattering experiments demonstrate that electrons are and remain point like. Even when we observe diffraction such as in a double-slit experiment where the wave-like nature is essential (interference!), we still detect point-like electrons when they hit the screen. Composite particles such as nuclei also do not dissolve.

How do we resolve this apparent contradiction between the existence of localized particles and the spreading of wave packets? A hint concerning the correct answer is given by the observation that we cannot predict where the next particle will hit the screen in the double-slit experiment, neither within a quantum-mechanical framework nor in any other way. Similarly, for the Stern–Gerlach experiment, we cannot predict if the next atom will be deflected up or down. This suggests a *statistical interpretation* of the wave function: $\psi(\vec{r}, t)$ *does not* describe how matter is distributed within space at the time t but rather with which *probability* the—intrinsically exactly or approximately point like—particle can be found at $\vec{r}$ at the time t.

Now we need to consider how $\psi(\vec{r}, t)$ exactly describes this probability. A reasonable first guess is that $\psi(\vec{r}, t)$ is the probability density, which meanings that $\psi(\vec{r}, t)\, d^3r$ is the probability to find the particle within the volume element d^3r. However, this cannot be correct because a probability density $\varrho(\vec{r}, t)$ must fulfill

1. $\varrho(\vec{r},t) \in \mathbb{R}$ und $\varrho(\vec{r},t) \geq 0 \,\forall\, \vec{r}, t$ and
2. $\int d^3r\, \varrho(\vec{r},t) = 1 \,\forall\, t$ (the full probability equals unity).

If we choose a function $\psi(\vec{r}, t=0)$ at a fixed time $t = 0$ with these properties, then the Schrödinger equation shows that the properties generally do not hold for $t > 0$: $\int d^3r\, \psi(\vec{r},t)$ is not conserved, nor does $\psi(\vec{r},t)$ remain positive semidefinite.

The ansatz

$$\varrho(\vec{r},t) = |\psi(\vec{r},t)|^2 = \psi^*(\vec{r},t)\,\psi(\vec{r},t), \tag{5.34}$$

however, doesn't lead to these problems. The first condition is automatically fulfilled. Furthermore, we can show that

$$\begin{aligned} \frac{d}{dt}\int d^3r\,|\psi(\vec{r},t)|^2 &= \int d^3r \left[\frac{\partial\psi^*}{\partial t}\psi + \psi^*\frac{\partial\psi}{\partial t}\right] \\ &= \int d^3r \left[\left(\frac{1}{i\hbar}\hat{H}\psi\right)^*\psi + \psi^*\frac{1}{i\hbar}\hat{H}\psi\right] \\ &= \frac{1}{i\hbar}\int d^3r \left[-(\hat{H}\psi)^*\,\psi + \psi^*\hat{H}\psi\right]. \end{aligned} \tag{5.35}$$

For the case of a single particle, we can express this as

$$\begin{aligned} (\hat{H}\psi)^* &= \left[-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r},t) + V(\vec{r})\psi(\vec{r},t)\right]^* \\ &= -\frac{\hbar^2}{2m}\nabla^2\psi^*(\vec{r},t) + V(\vec{r})\psi^*(\vec{r},t), \end{aligned} \tag{5.36}$$

and thus

$$\begin{aligned} \frac{d}{dt}\int d^3r\,|\psi(\vec{r},t)|^2 &= \frac{1}{i\hbar}\int d^3r \left[\frac{\hbar^2}{2m}\left(\nabla^2\psi^*\right)\psi - \cancel{V(\vec{r})\psi^*\psi} - \frac{\hbar^2}{2m}\,\psi^*\nabla^2\psi + \cancel{V(\vec{r})\psi^*\psi}\right] \\ &\overset{\text{by parts}}{=} \underbrace{\text{surface term}}_{=0} - \frac{1}{i\hbar}\frac{\hbar^2}{2m}\int d^3r\left[(\vec{\nabla}\psi^*)\cdot(\vec{\nabla}\psi) - (\vec{\nabla}\psi^*)\cdot(\vec{\nabla}\psi)\right] \\ &= 0. \end{aligned} \tag{5.37}$$

The surface terms disappear in our case if we assume that $\psi(\vec{r},t)$ decays quickly enough for $|\vec{r}| \to \infty$. Therefore, the full probability $\int d^3r\,|\psi|^2$ is conserved. So if we demand that $\int d^3r\,|\psi|^2 = 1$ in the initial state, then this normalization holds for all times. (Note that the ansatz $\rho = |\psi|$ would fulfill the first but not the second condition.)

Above, we have assumed a specific form of the Hamilton operator $\hat{H}$. Equation (5.35) shows that the full probability is conserved for a more general Hamilton operator, as long as

$$\int d^3r\,(\hat{H}\psi)^*\psi = \int d^3r\,\psi^*\hat{H}\psi \tag{5.38}$$

holds. This property of an operator $\hat{H}$ is called *hermiticity*. $\hat{H}$ must thus be *hermitian* for us to be allowed to interpret $|\psi|^2$ as a probability density. $\psi(\vec{r},t)$ itself is called the *probability amplitude*.

We can only consider functions $\psi(\vec{r},t)$ that fulfill the normalization condition

$$\int d^3r\,|\psi(\vec{r},t)|^2 = 1 \tag{5.39}$$

as possible wave functions for a single particle. More generally, it is sufficient to only demand that

$$\int d^3r\,|\psi(\vec{r},t)|^2 < \infty, \tag{5.40}$$

i.e., that the integral converges. If this is the case we can simply normalize $\psi(\vec{r},t)$ through multiplication with a number. Functions that fulfill Eq. (5.40) are called *square integrable* or *quadratically integrable*.

5.3.1 The Hilbert Space of Square-Integrable Functions

We will denote the set of square-integrable functions $f : M \to \mathbb{C}$ by $L_2(M)$. In single-particle quantum mechanics, we are especially interested in $L_2(\mathbb{R}^3)$. This set equipped with the usual addition of elements and multiplication with numbers from $\mathbb{C}$ is a *vector space* (linear space) over $\mathbb{C}$ because if $\psi_1(\vec{r}), \psi_2(\vec{r}) \in L_2(\mathbb{R}^3)$ (we will ignore the time argument for now since it is irrelevant in this context), then $\lambda_1\,\psi_1(\vec{r}) + \lambda_2\,\psi_2(\vec{r})$ is also square-integrable for any $\lambda_1, \lambda_2 \in \mathbb{C}$. This can be proved by

$$\begin{aligned}
\int d^3r\,|\lambda_1\,\psi_1(\vec{r}) + \lambda_2\,\psi_2(\vec{r})|^2 &\le \int d^3r\,(|\lambda_1\,\psi_1(\vec{r})| + |\lambda_2\,\psi_2(\vec{r})|)^2 \qquad \Big|\,\text{triangle inequality theorem} \\
&\le \int d^3r\,\left[(|\lambda_1\,\psi_1(\vec{r})| + |\lambda_2\,\psi_2(\vec{r})|)^2 + (|\lambda_1\,\psi_1(\vec{r})| - |\lambda_2\,\psi_2(\vec{r})|)^2\right] \\
&= 2\int d^3r\,\left[|\lambda_1|^2|\psi_1(\vec{r})|^2 + |\lambda_2|^2|\psi_2(\vec{r})|^2\right] \\
&= 2\,|\lambda_1|^2 \int d^3r\,|\psi_1(\vec{r})|^2 + 2\,|\lambda_2|^2 \int d^3r\,|\psi_2(\vec{r})|^2, \qquad (5.41)
\end{aligned}$$

which is finite because ψ_1 and ψ_2 are square-integrable. Note that at this point we need the definition of an integral according to *Lebesgue*, which is more universal then the integral according to Riemann. Loosely speaking, a function f is Lebesgue integrable if it only deviates from a Riemann-integrable (i.e., integrable in the usual sense) function $\tilde{f}$ on a null set, in which case the Lebesgue integral is equal to the Riemann integral of $\tilde{f}$. A null set is a set of measure zero, such as discrete points on a line. This definition is necessary for some proofs but we will not pay any further attention to it.

It is crucial for quantum mechanics that $L_2(\mathbb{R}^3)$ is not only a vector space but more specifically a *separable Hilbert space*. To understand the definition as well as further implications, we need a bit of preparation.

1. A *vector space* V over a field K fulfills the following axioms:
   (a) An addition $+ : V \otimes V \to V$ is defined on V and V is an abelian group with respect to this addition.
   (b) A multiplication with a scalar $\cdot : K \otimes V \to V$ exists (the operation symbol isn't explicitly written) and has the properties

$$\lambda\,(x + y) = \lambda\,x + \lambda\,y \qquad \text{(distributive)}, \tag{5.42}$$

$$(\lambda + \mu)\,x = \lambda\,x + \mu\,x \qquad \text{(distributive)}, \tag{5.43}$$

$$(\lambda\mu)\,x = \lambda\,(\mu x) \qquad \text{(associative)}, \tag{5.44}$$

$$1\,x = x, \tag{5.45}$$

   each for any λ, μ, x, y.

2. A *scalar product* between elements of a vector space V over a field $K = \mathbb{R}$ or $\mathbb{C}$ is a map $\langle \bullet, \bullet \rangle : V \otimes V \to K$ with the following properties:
   (a) $\langle x, x\rangle \in \mathbb{R}$ and $\langle x, x\rangle \ge 0$, where $\langle x, x\rangle = 0$ if and only if $x = 0$ (zero vector in V).
   (b) $\langle x, y\rangle = \langle y, x\rangle^*$, where * denotes complex conjugation (unnecessary for $K = \mathbb{R}$).
   (c)

$$\langle x, \lambda_1\,y_1 + \lambda_2\,y_2\rangle = \lambda_1\,\langle x, y_1\rangle + \lambda_2\,\langle x, y_2\rangle, \tag{5.46}$$

   i.e., the scalar product is linear in its second argument (in pure mathematics it is often defined as linear in the first argument, not the second). With (b) it follows that

$$\langle \lambda_1\,x_1 + \lambda_2\,x_2, y\rangle = \lambda_1^*\,\langle x_1, y\rangle + \lambda_2^*\,\langle x_2, y\rangle \tag{5.47}$$

   so that the scalar product in $K = \mathbb{C}$ is *anti-linear* (for $K = \mathbb{R}$ linear) in its first argument.

Two elements $x, y \in V$, $x, y \neq 0$, are called *orthogonal* if

$$\langle x, y\rangle = 0. \tag{5.48}$$

With the help of the scalar product, one can define a *norm*

$$\|x\| := \sqrt{\langle x, x\rangle} \tag{5.49}$$

on V.

3. A vector space with a scalar product is called a *unitary space*. (A vector space with a norm is called a normed vector space so that every unitary space also normed.) $L_2(\mathbb{R}^3)$ with the scalar product

$$\langle \varphi, \psi\rangle = \int d^3r\, \varphi^*(\vec{r})\, \psi(\vec{r}) \tag{5.50}$$

is a unitary space. To prove this, one must show that $\langle \varphi, \psi\rangle$ is actually a scalar product. This is quite simple but one must assume that two functions are seen as equal if they differ at most on a null set. For example, take

$$\psi(\vec{r}) = \begin{cases} 1 & \text{for} \quad \vec{r} \in \mathbb{Z}^3, \\ 0 & \text{otherwise.} \end{cases} \tag{5.51}$$

Then we know that

$$\langle \psi, \psi\rangle = \int d^3r\, \psi^2(\vec{r}) \overset{\text{Lebesgue!}}{=} 0. \tag{5.52}$$

The axiom $\langle \psi, \psi\rangle = 0 \Leftrightarrow \psi \equiv 0$ can only be fulfilled if we treat ψ as identical to the function

$$\tilde{\psi}(\vec{r}) = 0 \quad \forall\, \vec{r}. \tag{5.53}$$

The following concepts are only necessary because the unitary space $L_2(\mathbb{R}^3)$ is not of a finite dimension.

4. A series x_n of elements within a unitary (or at least normed) vector space is called a *Cauchy sequence* if for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ for which

$$\|x_m - x_n\| < \epsilon \quad \forall\, m, n > N, \tag{5.54}$$

i.e., the distances converge to zero for large indices.

5. A unitary (or normed) vector space V is called *complete* if every Cauchy sequence $x_n \in V$ converges strongly towards one element $x \in V$, or in other words, if for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ so that

$$\|x_n - x\| < \epsilon \quad \forall\, n > N. \tag{5.55}$$

This condition can also be written as

$$\lim_{n\to\infty} \|x_n - x\| = 0. \tag{5.56}$$

Every strongly converging series is also a Cauchy sequence but the reverse is not necessarily true, leading to the definition of completeness. It means that V doesn't have any "holes". For example, the set of rational numbers $\mathbb{Q}$ (which is a vector space over itself as a field but not a vector space over $\mathbb{R}$ or $\mathbb{C}$) is incomplete because the Cauchy sequence

$$\left\{ \left(1 + \frac{1}{n}\right)^n \,\middle|\, n \in \mathbb{N} \right\} = \left\{ 1, \left(\frac{3}{2}\right)^2, \left(\frac{4}{3}\right)^3, \left(\frac{5}{4}\right)^4, \ldots \right\} \tag{5.57}$$

within $\mathbb{Q}$ does not converge towards an element of $\mathbb{Q}$. The series converges in $\mathbb{R}$ with the limit e. It can be proven that an *finite-dimensional* unitary space is always complete.

6. A complete, unitary space is called a *Hilbert space*. (A complete, normed space is called a *Banach space*. It is a weaker definition, so every Hilbert space is also a Banach space but not vice versa.)

   $L_2(\mathbb{R}^3)$ is a Hilbert space. To prove this, we must show that it is complete. The Lebesgue integral is required for this step.

7. For quantum theory, we need yet another property: Not only do all Cauchy sequences have to converge, there must also be a series $F = \{x_n\}$ which comes arbitrarily close to every $x \in V$. One calls such a series *dense* in V. It can be shown that this is equivalent to saying that for every $x \in V$ there exists a partial series of F which converges towards x. If a dense series exists in V, i.e., if a dense countable subset of V exists, then V is called *separable*. Separability is an additional condition beyond the normal Hilbert-space properties. In physics textbooks, however, a Hilbert space is sometimes more strictly defined as a separable, complete unitary space.

   $L_2(\mathbb{R}^3)$ is a separable Hilbert space but we will forgo the proof of separability.

The separability of the Hilbert space V is important, as we will now discuss. First, let us define a *complete orthonormal system* in V as a set $M \subset V$ with the properties

1. $\|x\| = 1 \quad \forall\, x \in M$ (normalization),
2. $\langle x, y\rangle = 0 \quad \forall\, x, y \in M$ with $x \neq y$ (orthogonality),
3. there is no $v \in V$, $v \neq 0$, for which $\langle v, x\rangle = 0\ \forall\, x \in M$. One says that M is complete. Note that this is a different meaning of completeness than the completeness of the Hilbert space V.

Separability means that a series $F = \{f_1, f_2, \dots\}$ exists that comes arbitrarily close to every $x \in M$. F, as every series, is countable. We will now prove that a complete orthonormal system M is also countable.

Proof (*reductio ad absurdum*): We will assume that M is uncountable. For $x, y \in M,\ x \neq y$, we know that

$$\begin{aligned}\|x - y\| &= \sqrt{\langle x - y\,,\, x - y\rangle} \\ &= \sqrt{\langle x, x\rangle + \langle y, y\rangle - \underbrace{\langle x, y\rangle}_{=0} - \underbrace{\langle y, x\rangle}_{=0}} \\ &= \sqrt{\|x\|^2 + \|y\|^2} = \sqrt{2}. \end{aligned} \tag{5.58}$$

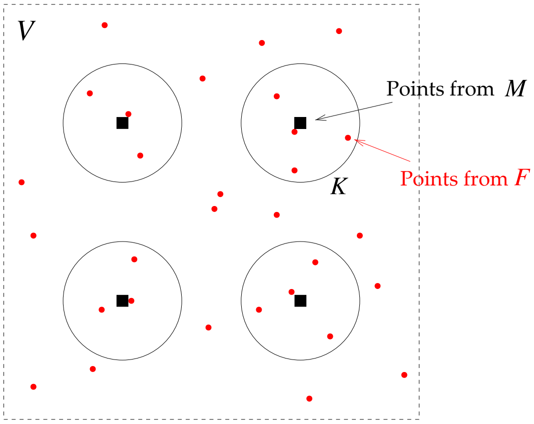
Two different elements of M thus always have a distance of $\sqrt{2}$. Now consider balls

$$K\Big(x, \frac{1}{2}\Big) := \Big\{z \in V \,\Big|\, \|z - x\| \leq \frac{1}{2}\Big\}, \tag{5.59}$$

centered around the elements $x \in M \subset V$. Since the radius of each ball is smaller than the distance between their centers, each ball $K(x, 1/2)$ contains exactly one element of M, namely x. Furthermore, the balls are disjoint because if there were a z with $z \in K(x, 1/2)$ and $z \in K(y, 1/2)$ with $x, y \in M$, $x \neq y$, then the distance between their centers would be

$$\begin{aligned}\|x - y\| &= \|(x - z) - (y - z)\| \\ &\leq \|x - z\| + \|-(y - z)\| = \|x - z\| + \|y - z\| \\ &= \frac{1}{2} + \frac{1}{2} = 1, \end{aligned} \tag{5.60}$$

using the triangle inequality. The result contradicts Eq. (5.58). Geometrically speaking, the radius of each ball is smaller than half the distance between two neighboring centers.



Since F is dense in V, each ball also contains at least one element $f \in F$ (actually, it contains a countable infinite number of them). There is exactly one ball for each $x \in M$ and they are disjoint. Then the cardinality of F is no less than the cardinality of M. But M is by assumption uncountable. Then F must also be uncountable and we obtain a contradiction.

The separability guaranties the existence of a countable orthonormal system. As a reminder, *finite-dimensional* Hilbert spaces are automatically complete and separable. For finite-dimensional Hilbert spaces, a finite orthonormal system always exists.

5.3.2 Expansion into Basis Functions

We saw that a separable Hilbert space, especially the space $L_2(\mathbb{R}^3)$, has complete, countable orthonormal systems. Each of these can be used as an orthonormal *basis* B of the Hilbert space, which means that one can expand any wave function $\psi \in L_2(\mathbb{R}^3)$ into basis elements $\varphi_n \in B \subset L_2(\mathbb{R})$. Since a separable Hilbert space has a countable basis, one calls its vector-space dimension countable.

It is useful to express the completeness of the basis through the *completeness relation*

$$\sum_n \varphi_n(\vec{r})\, \varphi_n^*(\vec{r}\,') = \delta(\vec{r} - \vec{r}\,'), \tag{5.61}$$

which we will now prove.

Since δ is actually a distribution, both sides of the relation are to be understood under an integral. The actual proposition is as follows: For all $f \in L_2(\mathbb{R}^3)$,

$$\int d^3r' \sum_n \varphi_n(\vec{r})\varphi_n^*(\vec{r}\,')f(\vec{r}\,') = \int d^3r'\, \delta(\vec{r} - \vec{r}\,')\, f(\vec{r}\,') \tag{5.62}$$

holds true. This is equivalent to

$$\sum_n \varphi_n(\vec{r}) \underbrace{\int d^3r'\, \varphi_n^*(\vec{r}\,')f(\vec{r}\,')}_{=\langle\varphi_n, f\rangle} = f(\vec{r}). \tag{5.63}$$

Because of the completeness of $L_2(\mathbb{R}^3)$ the sum (sequence) on the left-hand side converges. All that is left to show is that the series converges towards $f(\vec{r})$. We will do so again using *reductio ad absurdum*: Let us assume that the series converges towards a different function $g(\vec{r})$, so

$$\sum \varphi_n(\vec{r})\, \langle\varphi_n, f\rangle - f(\vec{r}) =: g(\vec{r}) \not\equiv 0. \tag{5.64}$$

Because of the completeness of $L_2(\mathbb{R}^3)$ under addition, g also lies in $L_2(\mathbb{R}^3)$. Then it follows that

$$\sum \underbrace{\langle\varphi_m, \varphi_n\rangle}_{=\delta_{m,n}} \langle\varphi_n, f\rangle - \langle\varphi_m, f\rangle = \langle\varphi_m, g\rangle \quad \forall\, \varphi_m \in B \tag{5.65}$$

$$\Rightarrow \quad \langle\varphi_m, f\rangle - \langle\varphi_m, f\rangle = 0 = \langle\varphi_m, g\rangle \quad \forall\varphi_m \in B. \tag{5.66}$$

But then g is orthogonal to all φ_m, and B is not a complete orthonormal system: contradiction!

Using the completeness relation we can easily expand any function $\varphi \in L_2(\mathbb{R}^3)$. The method uses insertion of the identity, which is very useful and will be employed quite often. It works like so:

$$\begin{aligned}
\psi(\vec{r}) &= \int d^3r'\, \delta(\vec{r}-\vec{r}\,')\, \psi(\vec{r}\,') \\
&= \int d^3r' \sum_n \varphi_n(\vec{r})\varphi_n^*(\vec{r}\,')\psi(\vec{r}\,') \\
&= \sum_n \varphi_n(\vec{r}) \underbrace{\int d^3r'\, \varphi_n^*(\vec{r}\,')\psi(\vec{r}\,')}_{=\langle\varphi_n,\psi\rangle} \\
&= \sum_n \langle\varphi_n, \psi\rangle\, \varphi_n(\vec{r}).
\end{aligned} \tag{5.67}$$

In this case, $\langle\varphi_n, \psi\rangle$ are (generally complex) numbers. They are the expansion coefficients of ψ in the basis $B = \{\varphi_n | n = 1, 2, \dots\}$.

5.3.3 Improper (Dirac) States

Often it is desirable to examine functions that aren't square integrable so that they aren't element of $L_2(\mathbb{R}^3)$. The Schrödinger equation also has sensible solutions that aren't in $L_2(\mathbb{R}^3)$. For example, the free Schrödinger equation

$$i\hbar\,\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\,\nabla^2\psi \tag{5.68}$$

is solved by plane waves

$$\psi = \psi_0\, e^{i(\vec{k}\cdot\vec{r}-\omega t)}. \tag{5.69}$$

These aren't square integrable:

$$\int d^3r\, |\psi|^2 = |\psi_0|^2 \int d^3r\, 1 = \infty. \tag{5.70}$$

If we extend the function space by plane waves and all possible superpositions, the result is another vector space. But its elements generally don't have a finite norm. Then we can't define Cauchy sequences and the space can't be complete (so it isn't a Hilbert space) and it isn't separable. So there is no guarantee that a countable basis exists. In fact, there is none. The functions $\psi = \psi_0\, e^{i\vec{k}\cdot\vec{r}}$ are linearly independent for different $\vec{k}$ but there are uncountably many $\vec{k}$, namely all $\vec{k} \in \mathbb{R}^3$. So the function space *can't* have any countable basis.

We will extend the space of allowed "functions" even further by including distributions such as $\delta(\vec{r}-\vec{R})$ with a fixed $\vec{R}$. Which mathematical objects are sensible is decided by the physics question. This "sensibly" extended space is called the *Dirac space*. One could also defend the point of view that physical states can only be described by wave functions in $L_2(\mathbb{R}^3)$ and that all other functions are only mathematical tools that simplify derivations and calculations. Since the observable part of the universe is also finite and the quantum-theoretical explanations shouldn't depend on the inobservable parts one can assume that the relevant Hilbert space should be countably infinite in dimension (see section **??**). Actually, some have argued that it has finite dimensions, albeit immense.

Typically, not square-integrable functions and distributions that are physically sensible at least allow a generalized orthonormality condition and a generalized completeness relation. A uncountable function system $\{\varphi_\alpha(\vec{r})\}$ with α from an uncountable set, e.g., $\alpha \in \mathbb{R}$, is called an orthonormal system if

$$\int d^3r\, \varphi_\alpha^*(\vec{r})\, \varphi_\beta(\vec{r}) = \delta(\alpha-\beta). \tag{5.71}$$

We furthermore call such a system complete if

$$\int d\alpha\, \varphi_\alpha(\vec{r})\, \varphi_\alpha^*(\vec{r}\,') = \delta(\vec{r}-\vec{r}\,'). \tag{5.72}$$

These conditions are fulfilled by plane waves:

Let $\varphi_{\vec{k}}(\vec{r}) = \frac{1}{(2\pi)^{3/2}}\, e^{i\vec{k}\cdot\vec{r}}$, then

$$\int d^3r\, \varphi^*_{\vec{k}}(\vec{r})\, \varphi_{\vec{k}'}(\vec{r}) = \frac{1}{(2\pi)^3}\, \underbrace{\int d^3r\, e^{-i\vec{k}\cdot\vec{r}+i\vec{k}'\cdot\vec{r}}}_{=\,(2\pi)^3\,\delta(\vec{k}-\vec{k}')} = \delta(\vec{k}-\vec{k}'). \tag{5.73}$$

This result is based on the identity

$$\int_{-\infty}^{\infty} dx\, e^{ikx} = 2\pi\,\delta(k), \tag{5.74}$$

which is valid "under an integral", which means that for every integrable test function $\tilde{f}(k)$ the relation

$$\int_{-\infty}^{\infty} dk\, \tilde{f}(k) \int_{-\infty}^{\infty} dx\, e^{ikx} = 2\pi \int_{-\infty}^{\infty} dk\, \tilde{f}(k)\, \delta(k) = 2\pi\, \tilde{f}(k=0) \tag{5.75}$$

holds. This relation can easily be understood: with suitable assumptions regarding convergence, the left-hand side reduces to

$$\begin{aligned}\int_{-\infty}^{\infty} dk\, \tilde{f}(k) \int_{-\infty}^{\infty} dx\, e^{ikx} &= 2\pi \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} \frac{dk}{2\pi}\, e^{ikx}\, \tilde{f}(k)\\ &= 2\pi \int_{-\infty}^{\infty} dx\, f(x) = 2\pi \int_{-\infty}^{\infty} dx\, e^{-i0x}\, f(x)\\ &= 2\pi\, \tilde{f}(k=0),\end{aligned} \tag{5.76}$$

where $f(x)$ is the (inverse) Fourier transform of $\tilde{f}(k)$.

Analogously to Eq. (5.73), we can derive that

$$\int d^3k\, \varphi_{\vec{k}}(\vec{r})\, \varphi^*_{\vec{k}}(\vec{r}\,') = \frac{1}{(2\pi)^3}\, \underbrace{\int d^3k\, e^{i\vec{k}\cdot\vec{r}-i\vec{k}\cdot\vec{r}\,'}}_{=\,(2\pi)^3\,\delta(\vec{r}-\vec{r}\,')} = \delta(\vec{r}-\vec{r}\,'). \tag{5.77}$$

Plane waves are thus complete in the generalized sense. Orthonormality and completeness allow expansion into a system of functions. For plane waves, this expansion is simply the well-known Fourier transformation.

δ distributions localized in position space also form an orthonormal system: Let $\varphi_{\vec{R}}(\vec{r}) = \delta(\vec{r}-\vec{R})$, then

$$\int d^3r\, \varphi^*_{\vec{R}}(\vec{r})\, \varphi_{\vec{R}'}(\vec{r}) = \int d^3r\, \delta(\vec{r}-\vec{R})\, \delta(\vec{r}-\vec{R}') = \delta(\vec{R}-\vec{R}'). \tag{5.78}$$

The orthonormal system is also complete, as shown by

$$\int d^3R\, \varphi_{\vec{R}}(\vec{r})\, \varphi^*_{\vec{R}}(\vec{r}\,') = \int d^3R\, \delta(\vec{r}-\vec{R})\, \delta(\vec{r}\,'-\vec{R}) = \delta(\vec{r}-\vec{r}\,'). \tag{5.79}$$

5.3.4 Local Conservation of Probability

We have demanded that the full probability $\int d^3r\, |\psi|^2$ stays conserved. This is a global conservation law. The dynamics as described by the Schrödinger equation leads to a more stringent conservation law: For a particle within a potential $V(\vec{r})$, we find that

$$\begin{aligned}\frac{\partial}{\partial t}|\psi(\vec{r},t)|^2 &= \left(\frac{\partial}{\partial t}\,\psi^*(\vec{r},t)\right)\psi(\vec{r},t) + \psi^*(\vec{r},t)\,\frac{\partial}{\partial t}\,\psi(\vec{r},t)\\ &= -\frac{1}{i\hbar}\,(\hat{H}\psi)^*\psi + \frac{1}{i\hbar}\,\psi^*\hat{H}\psi \qquad \Big|\ \text{since } \hat{H} \text{ is hermitian}\end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\,\nabla^2\psi^*\right)\psi - \cancel{\frac{1}{i\hbar}\left(V(\vec{r})\psi^*\right)\psi} + \frac{1}{i\hbar}\,\psi^*\left(-\frac{\hbar^2}{2m}\,\nabla^2\psi\right) + \cancel{\frac{1}{i\hbar}\,\psi^* V(\vec{r})\psi} \\
&= \frac{\hbar}{2mi}\left[(\nabla^2\psi^*)\psi - \psi^*\nabla^2\psi\right] \\
&= \frac{\hbar}{2mi}\left[\vec{\nabla}\cdot\left((\vec{\nabla}\psi^*)\psi\right) - (\vec{\nabla}\psi^*)\cdot(\vec{\nabla}\psi) - \vec{\nabla}\cdot\left(\psi^*\vec{\nabla}\psi\right) + (\vec{\nabla}\psi^*)\cdot(\vec{\nabla}\psi)\right] \\
&= -\vec{\nabla}\cdot\frac{\hbar}{2mi}\left[\psi^*\vec{\nabla}\psi - (\vec{\nabla}\psi^*)\psi\right], \qquad (5.80)
\end{aligned}$$

so basically

$$\frac{\partial}{\partial t}\,|\psi|^2 + \vec{\nabla}\cdot\frac{\hbar}{2mi}\left[\psi^*\vec{\nabla}\psi - (\vec{\nabla}\psi^*)\psi\right] = 0. \tag{5.81}$$

Since $\varrho = |\psi|^2$ is a type of *density* it makes sense to interpret the second term as the divergence of a *current density*

$$\vec{\jmath}(\vec{r},t) := \frac{\hbar}{2mi}\left[\psi^*(\vec{r},t)\vec{\nabla}\psi(\vec{r},t) - \left(\vec{\nabla}\psi^*(\vec{r},t)\right)\psi(\vec{r},t)\right]. \tag{5.82}$$

$\vec{\jmath}$ is called the *probability current density.* This yields a *continuity equation*

$$\frac{\partial\varrho}{\partial t} + \vec{\nabla}\cdot\vec{\jmath} = 0 \tag{5.83}$$

similar to the one of electrodynamics but for probability instead of charge. It describes a type of *local* conservation of probability: ϱ can only be changed by probability flowing in or out.

5.3.5 Momentum Representation

Fourier transforms are defined for square-integrable functions and for the physically relevant improper states. In quantum mechanics, we often express it as a function of the momentum $\vec{p} = \hbar\vec{k}$ instead of the wave vector $\vec{k}$:

$$\tilde{\psi}(\vec{p},t) = \int d^3r\, e^{-i\vec{p}\cdot\vec{r}/\hbar}\,\psi(\vec{r},t), \tag{5.84}$$

$$\psi(\vec{r},t) = \int \frac{d^3p}{(2\pi\hbar)^3}\, e^{i\vec{p}\cdot\vec{r}/\hbar}\,\tilde{\psi}(\vec{p},t) = \int \frac{d^3p}{h^3}\, e^{i\vec{p}\cdot\vec{r}/\hbar}\,\tilde{\psi}(\vec{p},t). \tag{5.85}$$

This asymmetrical definition of the normal and reverse transformation is common in physics. ψ and $\tilde{\psi}$ carry the same information. $\tilde{\psi}(\vec{p},t)$ is called *the wave function in momentum space.* The concept of equivalent representations in position space and in momentum space is rooted in classical Hamilton mechanics, in which generalized coordinates and generalized momenta essentially appear on equal footing. If $\psi(\vec{r},t)$ is normalized,

$$\int d^3r\,|\psi(\vec{r},t)|^2 = 1, \tag{5.86}$$

then Plancherel's theorem gives

$$\begin{aligned}
\int \frac{d^3p}{(2\pi\hbar)^3}\,|\tilde{\psi}(\vec{p},t)|^2 &= \int \frac{d^3p}{(2\pi\hbar)^3}\,\tilde{\psi}^*(\vec{p},t)\,\tilde{\psi}(\vec{p},t) \\
&= \int \frac{d^3p}{(2\pi\hbar)^3}\int d^3r\, e^{i\vec{p}\cdot\vec{r}/\hbar}\,\psi^*(\vec{r},t)\int d^3r'\, e^{-i\vec{p}\cdot\vec{r}\,'/\hbar}\,\psi(\vec{r}\,',t) \\
&= \int d^3r\, d^3r'\,\psi^*(\vec{r},t)\,\psi(\vec{r}\,',t)\underbrace{\int \frac{d^3p}{(2\pi\hbar)^3}\, e^{i\vec{p}\cdot(\vec{r}-\vec{r}\,')/\hbar}}_{\substack{=\frac{1}{(2\pi\hbar)^3}\,(2\pi)^3\,\delta\left(\frac{\vec{r}-\vec{r}\,'}{\hbar}\right) \\ =\delta(\vec{r}-\vec{r}\,')}} \\
&= \int d^3r\,|\psi(\vec{r},t)|^2 = 1. \qquad (5.87)
\end{aligned}$$

Hence, the normalization is conserved under Fourier transformation. Since $\tilde{\psi}$ is normalized and $|\tilde{\psi}|$ is obviously positive, interpreting $\tilde{\psi}$ as a probability density makes perfect sense.

We can also rewrite the Schrödinger equation in momentum space, giving

$$i\hbar\,\frac{\partial}{\partial t}\tilde{\psi}(\vec{p},t) = \int d^3r\, e^{-i\vec{p}\cdot\vec{r}/\hbar}\, i\hbar\,\frac{\partial}{\partial t}\,\psi(\vec{r},t) = \int d^3r\, e^{-i\vec{p}\cdot\vec{r}/\hbar}\Big[-\frac{\hbar^2}{2m}\,\nabla^2\psi(\vec{r},t) + V(\vec{r})\,\psi(\vec{r},t)\Big]. \tag{5.88}$$

By double integration by parts in the kinetic-energy term, we find that

$$\begin{aligned}\ldots &= \int d^3r\,\Big[-\frac{\hbar^2}{2m}\Big(-\frac{i\vec{p}}{\hbar}\Big)^2 e^{-i\vec{p}\cdot\vec{r}/\hbar}\,\psi(\vec{r},t) + V(\vec{r})\,e^{-i\vec{p}\cdot\vec{r}/\hbar}\,\psi(\vec{r},t)\Big]\\ &= \int d^3r\,\Big[\frac{p^2}{2m}\,e^{-i\vec{p}\cdot\vec{r}/\hbar}\,\psi(\vec{r},t) + V(i\hbar\,\vec{\nabla}_{\vec{p}})\,e^{-i\vec{p}\cdot\vec{r}/\hbar}\,\psi(\vec{r},t)\Big]. \end{aligned} \tag{5.89}$$

The last identity is best understood if we expand $V(\vec{r})$ into a power series. Each power of $i\hbar\,\vec{\nabla}_{\vec{p}}$ extracts the exact same power of $\vec{r}$ from the exponent. Now we can integrate over the series, resulting in

$$i\hbar\,\frac{\partial}{\partial t}\,\tilde{\psi}(\vec{p},t) = \frac{p^2}{2m}\,\tilde{\psi}(\vec{p},t) + V(i\hbar\,\vec{\nabla}_{\vec{p}})\,\tilde{\psi}(\vec{p},t). \tag{5.90}$$

In momentum space, the momentum operator is simply a multiplication with a number vector $\vec{p}$,

$$\hat{\vec{p}} = \vec{p}. \tag{5.91}$$

On the other hand, the position operator is now more complex, mainly consisting of a gradient with respect to the momentum,

$$\hat{\vec{r}} = i\hbar\,\vec{\nabla}_{\vec{p}}. \tag{5.92}$$

There is a clear analogy between the position operator in momentum space and the momentum operator in position space, $\hat{\vec{p}} = -i\hbar\,\vec{\nabla}$.

For certain potentials, using momentum space can simplify the solution of the Schrödinger equation. For example, the free particle equation is actually turned into the *ordinary* (not partial) differential equation

$$i\hbar\,\frac{d\tilde{\psi}}{dt} = \frac{p^2}{2m}\,\tilde{\psi}. \tag{5.93}$$

Furthermore, momentum space is often utilized in condensed matter physics since it is very useful for periodic lattice potentials (such as crystal lattices). Moreover, it can also simplify certain formal derivations, as we will also see in the following section.

5.4 Expectation Values and Deviations

Since we have identified $|\psi(\vec{r},t)|^2$ as a probability density, we can immediately specify expressions for the expectation values of purely position-dependent quantities $A(\vec{r})$. We denote expectation values with angle brackets,

$$\langle A(\vec{r})\rangle = \int d^3r\,|\psi(\vec{r},t)|^2\,A(\vec{r}) = \int d^3r\,\psi^*(\vec{r},t)\,A(\vec{r})\,\psi(\vec{r},t). \tag{5.94}$$

For purely momentum-dependent quantities, we switch to momentum space,

$$\langle B(\vec{p})\rangle = \int \frac{d^3p}{(2\pi\hbar)^3}\,|\tilde{\psi}(\vec{p},t)|^2\,B(\vec{p}) = \int \frac{d^3p}{(2\pi\hbar)^3}\,\tilde{\psi}^*(\vec{p},t)\,B(\vec{p})\,\tilde{\psi}(\vec{p},t). \tag{5.95}$$

The expectation value $\langle\vec{p}\rangle$ was already calculated in section 4.2. With only minimally changed notation, it is

$$\langle\vec{p}\rangle = \int \frac{d^3p}{(2\pi\hbar)^3}\,\tilde{\psi}^*(\vec{p},t)\,\vec{p}\,\tilde{\psi}(\vec{p},t)$$

$$= \int d^3r\, d^3r' \,\frac{d^3p}{(2\pi\hbar)^3}\, e^{i\vec{p}\cdot\vec{r}/\hbar}\, \psi^*(\vec{r},t)\, \vec{p}\, e^{-i\vec{p}\cdot\vec{r}\,'/\hbar}\, \psi(\vec{r}\,',t)$$
$$= \int d^3r\, d^3r' \,\frac{d^3p}{(2\pi\hbar)^3}\, e^{i\vec{p}\cdot\vec{r}/\hbar}\, \psi^*(\vec{r},t) \left(-\frac{\hbar}{i}\right) \vec{\nabla}' e^{-i\vec{p}\cdot\vec{r}\,'/\hbar}\, \psi(\vec{r}\,',t). \tag{5.96}$$

The main idea in the derivation was to express the momentum $\vec{p}$ in terms of a gradient of the exponential function. This probably becomes more clear when reading the chain of equations backwards. Using integration by parts in $\vec{r}\,'$, we now move the gradient to the other factor dependent on $\vec{r}\,'$,

$$\begin{aligned}\langle \vec{p}\rangle &= \int d^3r\, d^3r' \,\frac{d^3p}{(2\pi\hbar)^3}\, e^{i\vec{p}\cdot\vec{r}/\hbar}\, \psi^*(\vec{r},t) e^{-i\vec{p}\cdot\vec{r}\,'/\hbar}\, \frac{\hbar}{i}\, \vec{\nabla}'\psi(\vec{r}\,',t)\\
&= \int d^3r\, d^3r'\, \psi^*(\vec{r},t) \left[\frac{\hbar}{i}\,\vec{\nabla}'\psi(\vec{r}\,',t)\right] \underbrace{\int \frac{d^3p}{(2\pi\hbar)^3}\, e^{i\vec{p}\cdot(\vec{r}-\vec{r}\,')/\hbar}}_{\delta(\vec{r}-\vec{r}\,')}\\
&= \int d^3r\, \psi^*(\vec{r},t)\, \frac{\hbar}{i}\, \vec{\nabla}\psi(\vec{r},t)\\
&= \int d^3r\, \psi^*(\vec{r},t)\, \hat{\vec{p}}\psi(\vec{r},t),\end{aligned} \tag{5.97}$$

with the momentum operator $\hat{\vec{p}} = -i\hbar\,\vec{\nabla}$ in position space.

Analogously, we find the more general relation

$$\langle B(\vec{p})\rangle = \int d^3r\, \psi^*(\vec{r},t)\, B\left(\frac{\hbar}{i}\,\vec{\nabla}\right) \psi(\vec{r},t) = \int d^3r\, \psi^*(\vec{r},t)\, B(\hat{\vec{p}})\, \psi(\vec{r},t). \tag{5.98}$$

In full generality, we write for any observable $C(\vec{r},\vec{p})$,

$$\langle C\rangle = \int d^3r\, \psi^*(\vec{r},t)\, C\Big(\vec{r}, \frac{\hbar}{i}\,\vec{\nabla}\Big)\, \psi(\vec{r},t). \tag{5.99}$$

Note that the order of the factors r_i and p_i in C is relevant, unlike in classical physics, see section 5.1.

We can also express the last expectation value in momentum space,

$$\langle C\rangle = \int \frac{d^3p}{(2\pi\hbar)^3}\, \tilde{\psi}^*(\vec{p},t)\, C\big(i\hbar\vec{\nabla}_{\vec{p}}, \vec{p}\big)\, \tilde{\psi}(\vec{p},t). \tag{5.100}$$

5.4.1 Hermiticity

Since measurable quantities must have real values—we don't even know what it would mean to measure a complex quantity—it is also sensible to demand that the expectation values of measurable variables must be real. Which properties must an operator $\hat{C}$ representing a measurable variable posses to guarantee this property? We have

$$\big\langle \hat{C}\big\rangle = \int d^3r\, \psi^*(\vec{r},t)\, \hat{C}\, \psi(\vec{r},t), \tag{5.101}$$

which also means that

$$\big\langle \hat{C}\big\rangle^* = \int d^3r\, \psi(\vec{r},t)\, \big[\hat{C}\, \psi(\vec{r},t)\big]^* = \int d^3r\, \big[\hat{C}\, \psi(\vec{r},t)\big]^*\, \psi(\vec{r},t). \tag{5.102}$$

Demanding $\big\langle \hat{C}\big\rangle^* = \big\langle \hat{C}\big\rangle$ leads to

$$\int d^3r\, \big[\hat{C}\, \psi(\vec{r},t)\big]^*\, \psi(\vec{r},t) = \int d^3r\, \psi^*(\vec{r},t)\, \hat{C}\, \psi(\vec{r},t), \tag{5.103}$$

which must hold for all square-integrable functions $\psi(\vec{r},t)$. If this condition be fulfilled the operator $\hat{C}$ is called *hermitian*, as we previously discussed for the Hamilton operator as a special case.

We consider the momentum operator $\langle p \rangle = \frac{\hbar}{i}\vec{\nabla}$ as an example. We have

$$\begin{aligned}
\int d^3r \left[\frac{\hbar}{i}\vec{\nabla}\psi(\vec{r},t)\right]^* \psi(\vec{r},t) &= \int d^3r \left(-\frac{\hbar}{i}\right)\left[\vec{\nabla}\psi^*(\vec{r},t)\right]\psi(\vec{r},t) \\
&\overset{\text{by parts}}{=} \int d^3r \,\frac{\hbar}{i}\,\psi^*(\vec{r},t)\,\vec{\nabla}\psi(\vec{r},t) \\
&= \int d^3r\,\psi^*(\vec{r},t)\,\frac{\hbar}{i}\,\vec{\nabla}\psi(\vec{r},t), \qquad (5.104)
\end{aligned}$$

which proves the momentum operator is hermitian. Hermiticity can also easily be shown for the Hamilton operator $\hat{H} = \hat{p}^2/2m + V(\vec{r})$.

5.4.2 Deviations

Similar to expectation values, higher moments of the distribution can also be calculated, in particular deviations. We will see that in quantum mechanics, a nonzero deviation or scattering of measurement values must exists for *fundamental* reasons, even for ideal experimental circumstances that do not introduce measurement errors. We define the *root mean square deviation* of a quantity $C(\vec{r},\vec{p})$ as

$$\begin{aligned}
\Delta C &:= \sqrt{\langle (C - \langle C \rangle)^2 \rangle} = \sqrt{\langle C^2 - 2C\langle C \rangle + \langle C \rangle^2 \rangle} \\
&= \sqrt{\langle C^2 \rangle - 2\langle C \rangle\langle C \rangle + \langle C \rangle^2} \\
&= \sqrt{\langle C^2 \rangle - \langle C \rangle^2}. \qquad (5.105)
\end{aligned}$$

ΔC^2 is called the *mean square deviation.*

Note that $\langle C^2 \rangle$ is not the same as $\langle C \rangle^2$. For example, if C assumes values of ± 1 with the same probability, then $\langle C^2 \rangle = \langle 1 \rangle = 1$, but

$$\langle C \rangle^2 = \left(\frac{1}{2} \times 1 + \frac{1}{2} \times (-1)\right)^2 = 0^2 = 0. \qquad (5.106)$$

Another example: consider a Gaussian wave packet in one dimension. Let us take the wave function in position space

$$\psi(x, t = 0) = \frac{1}{(2\pi)^{1/4}\sqrt{\sigma}}\, e^{ip_0x/\hbar}\, \exp\left(-\frac{(x-x_0)^2}{4\sigma^2}\right). \qquad (5.107)$$

We aren't interested in how it develops over time, so we will choose a fixed time $t = 0$. That leaves

$$|\psi(x)|^2 = \frac{1}{\sqrt{2\pi}\sigma}\, \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right), \qquad (5.108)$$

which means that the probability density is a normalized Gaussian around the average value x_0 with a width of σ. The mean square deviation of the position is

$$\begin{aligned}
\Delta x^2 = (\Delta x)^2 = \langle (x - \langle x \rangle)^2 \rangle &= \langle (x - x_0)^2 \rangle \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} dx\,(x-x_0)^2\, \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right) \\
&\overset{u = x - x_0}{=} \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} du\, u^2\, \exp\left(-\frac{u^2}{2\sigma^2}\right) = \sigma^2. \qquad (5.109)
\end{aligned}$$

We can calculate the momentum deviation using the momentum operator in position space or the wave function in momentum space. We will do the latter, resulting in

$$\begin{aligned}
\tilde{\psi}(p) &= \int dx\, e^{-ipx/\hbar}\,\psi(x) \\
&= \frac{1}{(2\pi)^{(1/4)}\sqrt{\sigma}} \int dx\, e^{-i(p-p_0)x/\hbar}\, \exp\left(-\frac{(x-x_0)^2}{4\,\sigma^2}\right)
\end{aligned}$$

$$= 2^{3/4}\pi^{1/4}\sqrt{\sigma}\, e^{-i(p-p_0)x_0/\hbar}\, \exp\left(-\frac{\sigma^2\,(p-p_0)^2}{\hbar^2}\right). \tag{5.110}$$

The Fourier transform of a normalized Gaussian is another normalized Gaussian. The momentum's mean value is obviously p_0. The mean square deviation is

$$\begin{aligned} \Delta p^2 &= \left\langle (p-p_0)^2 \right\rangle \\ &= 2^{3/2}\sqrt{\pi}\,\sigma \int \frac{dp}{2\pi\hbar}\,(p-p_0)^2\,\exp\left(-\frac{2\,\sigma^2\,(p-p_0)^2}{\hbar^2}\right) \\ &\overset{p'=p-p_0}{=} \sqrt{\frac{2}{\pi}}\,\frac{\sigma}{\hbar}\int dp'\,(p')^2\,\exp\left(-\frac{2\,\sigma^2\,(p')^2}{\hbar^2}\right) = \frac{\hbar^2}{4\,\sigma^2}. \end{aligned} \tag{5.111}$$

This shows that the Fourier transform $\tilde{\psi}(p)$ becomes wider the narrower the original function $\psi(x)$ is. This is a general characteristic of the Fourier transformation. More specifically, we find that

$$\Delta x\,\Delta p = \sigma\,\frac{\hbar}{2\,\sigma} = \frac{\hbar}{2}. \tag{5.112}$$

The specific value $\hbar/2$ is based on our choice of a Gaussian for $\psi(x)$.

5.5 Uncertainty Relation of Position and Momentum

In the last section, we found $\Delta x\Delta p = \hbar/2$ for Gaussian wave packets. We would now like to generalize the solution to any wave function. The deduction is mathematically somewhat elaborate and starts with the definition of an auxiliary function

$$\begin{aligned} I(\lambda) &:= \int_{-\infty}^{\infty} dx\ \left|(x-\langle x\rangle)\psi(x) + i\lambda(\hat{p}-\langle\hat{p}\rangle)\psi(x)\right|^2 \\ &= \int_{-\infty}^{\infty} dx\ \left|(x-\langle x\rangle)\psi(x) + i\lambda\left(\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle\hat{p}\rangle\right)\psi(x)\right|^2. \end{aligned} \tag{5.113}$$

Since the integrand is an absolute value squared $I(\lambda) \geq 0$ hold for all λ. We rewrite the integral as

$$\begin{aligned} I(\lambda) &= \int_{-\infty}^{\infty} dx\,\psi^*(x)(x-\langle x\rangle)^2\psi(x) \\ &\quad + \int_{-\infty}^{\infty} dx\,\psi^*(x)(x-\langle x\rangle)i\lambda\left(\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle\hat{p}\rangle\right)\psi(x) \\ &\quad - \int_{-\infty}^{\infty} dx\,i\lambda\left[\left(-\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle\hat{p}\rangle\right)\psi^*(x)\right](x-\langle x\rangle)\psi(x) \\ &\quad + \int_{-\infty}^{\infty} dx\,\lambda\left[\left(-\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle\hat{p}\rangle\right)\psi^*(x)\right]\lambda\left(\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle\hat{p}\rangle\right)\psi(x) \\ &\overset{\text{by parts}}{=} \Delta x^2 + i\lambda\int_{-\infty}^{\infty} dx\,\psi^*(x)(x-\langle x\rangle)\left(\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle\hat{p}\rangle\right)\psi(x) \\ &\quad - i\lambda\int_{-\infty}^{\infty} dx\,\psi^*(x)\left(\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle\hat{p}\rangle\right)(x-\langle x\rangle)\psi(x) \\ &\quad + \lambda^2\int_{-\infty}^{\infty} dx\,\psi^*(x)\left(\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle\hat{p}\rangle\right)^2\psi(x) \\ &= \Delta x^2 + i\lambda\cancel{\int_{-\infty}^{\infty} dx\,\psi^*(x)(x-\langle x\rangle)\left(\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle\hat{p}\rangle\right)\psi(x)} \\ &\quad - i\lambda\int_{-\infty}^{\infty} dx\,\psi^*(x)\frac{\hbar}{i}\psi(x) \end{aligned}$$

$$
\begin{aligned}
&\cancel{- i\lambda \int_{-\infty}^{\infty} dx\, \psi^*(x)(x - \langle x\rangle)\left(\frac{\hbar}{i}\frac{\partial}{\partial x} - \langle \hat{p}\rangle\right)\psi(x)} \\
&+ \lambda^2 \Delta p^2 \\
= {}& \Delta x^2 - \hbar\lambda + \Delta p^2\, \lambda^2. \qquad (5.114)
\end{aligned}
$$

It follows that

$$\Delta x^2 \geq -\Delta p^2 \lambda^2 + \hbar\lambda \quad \forall \lambda. \tag{5.115}$$

Since this is true for all λ it must in particular hold for the λ that maximizes the right-hand side. This value for λ can be calculated from

$$\frac{d}{d\lambda}(-\Delta p^2 \lambda^2 + \hbar\lambda) = -2\Delta p^2 \lambda + \hbar = 0 \tag{5.116}$$

$$\Rightarrow \quad \lambda = \frac{\hbar}{2\Delta p^2}. \tag{5.117}$$

Hence, we conclude

$$\Delta x^2 \geq -\Delta p^2 \frac{\hbar^2}{4\Delta p^4} + \frac{\hbar^2}{2\Delta p^2} = \frac{\hbar^2}{4\Delta p^2} \tag{5.118}$$

as well as

$$\Delta x^2 \Delta p^2 \geq \frac{\hbar^2}{4} \tag{5.119}$$

and, finally,

$$\Delta x\, \Delta p \geq \frac{\hbar}{2}. \tag{5.120}$$

This is called the *Heisenberg position-momentum uncertainty relation* (or uncertainly principle). This relation tells us that it is fundamentally impossible to exactly measure both the position and the momentum of a particle. The reason is that measuring both $\Delta x = 0$ and $\Delta p = 0$ would contradict the uncertainty relation. The appearance of $\hbar$ demonstrates that the relation is a consequence of quantum mechanics—the classical limit corresponds to $\hbar \to 0$, after all.

If we look at the derivation we can see that the result stems from the appearance of the term $-\hbar\lambda$ in Eq. (5.114). This term came from the difference of

$$\int_{-\infty}^{\infty} dx\, \psi^*(x)\,(x - \langle x\rangle)\,(\hat{p} - \langle \hat{p}\rangle)\,\psi(x) \tag{5.121}$$

and

$$\int_{-\infty}^{\infty} dx\, \psi^*(x)\,(\hat{p} - \langle \hat{p}\rangle)\,(x - \langle x\rangle)\,\psi(x), \tag{5.122}$$

and thus from the non-commutativity of position and momentum, which we saw in 5.1:

$$[\hat{x}, \hat{p}] = i\hbar. \tag{5.123}$$

We will get back to the connection between commutators and the uncertainty principles in section **??**. There, a different and more elegant, operator-based, derivation will be given that doesn't rely on wave functions.

5.6 Measurements I

The Schrödinger theory developed until this point is of *statistical* nature: it provides statements about probability distributions. The distribution functions cannot be arbitrarily exact for all measurable quantities, as the position-momentum uncertainty relation demonstrates. If we know the exact position of a particle, for example, then we have no information about its momentum. This statistical character is unavoidable within the standard quantum-theoretical perspective. It raises the question whether this uncertainty is actually a property of the real world or a weakness of standard quantum theory. In other words, which of the following statements is true?

- The wave function describes the position and momentum of a particle as exactly as possible. Thus, the particle does not have any properties concerning position and momentum that cannot be described via its wave function. In particular, it does not have an exact position and an exact momentum at the same time.
- In reality, a particle has an exact position and exact momentum at any given time. The wave function thus only provides an incomplete description of the particle. Since standard quantum theory cannot describe a particle with an exact position and exact momentum as a matter of principle, although they really exist, quantum theory must be incomplete.

This question is closely related to the interpretation of the *measurement process*. What happens during a measurement of an observable A in a system with a given wave function ψ? According to the statistical interpretation, $|\psi(\vec{r},t)|^2$ is the probability density of a particle in position space at the time t. Let's assume that the wave function isn't perfectly localized before the measurement. At the time $t=0$ we then measure, say, the position of a particle with a perfect measuring device. Directly before the measurement, the probability density is $|\psi(\vec{r},0)|^2$. Directly afterwards, i.e., at the time $t=dt$, we know where the particle is. Let's say we observe the particle at the position $\vec{R}$. Then the probability density at the time $t=dt$ isn't $|\psi(\vec{r},0)|^2$ anymore since $|\psi(\vec{r},0)|^2$ expresses that the particle can also be at positions $\vec{r}\neq\vec{R}$, which isn't the case. We can conclude that the probability density, and with it the wave function $\psi(\vec{r},t)$, has to change discontinuously (i.e., jump) when the measurement takes place. This isn't an artifact of the idealized exact measurement. If we interpret $|\psi(\vec{r},t)|^2$ as a probability density that expresses our knowledge of the system, then any change of that knowledge, especially any successful measurement, must lead to a change of the wave function.

This conclusion leads back to the original question: Does the measurement lead to the localization at point $\vec{R}$ of the previously delocalized particle or does it only reveal the particle's position $\vec{R}$, which existed already before the measurement but was unknown? From the perspective of classical physics, the second explanation is surely the more natural of the two. But it actually isn't tenable, as we have yet to see.

Now we will consider a measurement of any arbitrary observable $\hat{A}$. We denote it with a circumflex since observables are represented by operators. (Position measurements turn out to be unfavorable for our discussion of measurement processes because wave functions with exact positions are δ distributions and thus not square integrable.) Let the wave function be $\psi(\vec{r},0)$ immediately before the measurement. Performing the measurement of $\hat{A}$, we find an exact value A. (We here only consider measurements that exactly determine the measured value, not only restrict it.) Due to the measurement, the wave function must have changed in such a way that the observable $\hat{A}$ now has the exact value A. This means that the expectation value $\langle\hat{A}\rangle$ has to coincide with A and the deviation ΔA has to vanish.

The expectation value of the observable is (we suppress the time argument $t=dt$)

$$\langle\hat{A}\rangle = \int d^3r\,\psi^*(\vec{r})\,\hat{A}\,\psi(\vec{r}) \overset{!}{=} A. \tag{5.124}$$

The mean square deviation is

$$\Delta A^2 = \int d^3r\,\psi^*(\vec{r})\,(\hat{A}-A)^2\,\psi(\vec{r}). \tag{5.125}$$

Being an observable, $\hat{A}$ must be hermitian. Then $\hat{A}-A$ must also be hermitian, since multiplication with the real constant A is also hermitian. This gives us

$$\Delta A^2 = \int d^3r\,\left[(\hat{A}-A)\,\psi(\vec{r})\right]^*(\hat{A}-A)\,\psi(\vec{r}) = \|(\hat{A}-A)\,\psi(\vec{r})\|^2 \overset{!}{=} 0. \tag{5.126}$$

The last expression is positive semidefinite and it is zero if and only if

$$(\hat{A}-A)\,\psi(\vec{r}) = 0 \quad \forall\vec{r}, \tag{5.127}$$

i.e., if

$$\hat{A}\,\psi(\vec{r}) = A\,\psi(\vec{r}) \quad \forall\vec{r}. \tag{5.128}$$

Hence, the wave function must fulfill this equation immediately after measuring the value A.

The derived equation for deviation-free wave functions,

$$\hat{A}\,\psi_A(\vec{r}) = A\,\psi_A(\vec{r}), \tag{5.129}$$

is an *eigenvalue equation* for the linear operator $\hat{A}$ on $L_2(\mathbb{R}^3)$. A is the *eigenvalue* to the *eigenfunction* $\psi_A(\vec{r})$. These definitions are closely related to the eigenvectors and eigenvalues of matrices. We denote the eigenvalue A as a subscript of ψ_A to indicate that this equation is only valid for the eigenfunction with eigenvalue A and not for arbitrary wave functions. A measurement can only provide values that permit a zero deviation $\Delta A = 0$, i.e., it can only give eigenvalues. After the measurement, the wave function is an eigenfunction of $\hat{A}$ to the measured eigenvalue.

To sum up, we have found the following:

- When measuring an observable $\hat{A}$, the wave function suddenly and discontinuously transitions into an eigenfunction of $\hat{A}$ (the wave function "collapses").
- The measured value is the eigenvalue to that eigenfunction.

These statements constitute the *von Neumann projection postulate.* It is actually an additional postulate since it does not follow from the time evolution as described by the Schrödinger equation, which we shall see.

Since $L_2(\mathbb{R}^3)$ is a separable Hilbert space a countable orthonormal basis exists. Actually, infinitely many such bases exist. One can prove that the eigenfunctions of any arbitrary hermitian operator $\hat{A}$ on $L_2(\mathbb{R}^3)$ form such a basis. (For now, we will consider the case that the eigenvalues are not *degenerate*, i.e., that a single A only corresponds to a single eigenfunction ψ_A.) This has important consequences.

Let us consider the expectation value of $\hat{A}$ once more,

$$\langle \hat{A} \rangle = \int d^3r\,\psi^*(\vec{r})\,\hat{A}\,\psi(\vec{r}). \tag{5.130}$$

The completeness of the basis of eigenfunctions implies that

$$\sum_A \psi_A(\vec{r})\,\psi_A^*(\vec{r}\,') = \delta(\vec{r}-\vec{r}\,'). \tag{5.131}$$

Inserting this completeness relation into Eq. (5.130) leads to

$$\begin{aligned}
\langle \hat{A} \rangle &= \int d^3r\,d^3r'\,\psi^*(\vec{r})\,\hat{A}\,\delta(\vec{r}-\vec{r}\,')\,\psi(\vec{r}\,') \\
&= \int d^3r\,d^3r'\,\psi^*(\vec{r})\,\hat{A}\sum_A \psi_A(\vec{r})\psi_A^*(\vec{r}\,')\psi(\vec{r}\,') && \Big|\ \hat{A}\text{ only acts on }\vec{r} \\
&= \sum_A \int d^3r\,d^3r'\,\psi^*(\vec{r})\,A\,\psi_A(\vec{r})\psi_A^*(\vec{r}\,')\psi(\vec{r}\,') && \Big|\ \text{with Eq. (5.129)} \\
&= \sum_A A \int d^3r\,\psi^*(\vec{r})\psi_A(\vec{r}) \int d^3r'\,\psi_A^*(\vec{r}\,')\psi(\vec{r}\,') \\
&= \sum_A A\,\langle \psi, \psi_A\rangle\langle \psi_A, \psi\rangle \\
&= \sum_A \big|\langle \psi_A, \psi\rangle\big|^2 A.
\end{aligned} \tag{5.132}$$

So the expectation value is the weighted average of the eigenvalues A with the weights $|\langle \psi_A, \psi\rangle|^2$.

What is the meaning of the expectation value for measurements? Since measured values can only be eigenvalues and the expectation value doesn't have to be one (it generally lies somewhere between two discrete eigenvalues) the expectation value can't generally describe the "expected result" of a measurement. If we take the interpretation of $|\psi|^2$ as a distribution function seriously and if Eq. (5.130) is correct for determining the eigenvalues, then $\langle \hat{A} \rangle$

should represent the averaged value over multiple measurements for equally prepared systems. So, if we take N measurements of $\hat{A}$ and measure values A_n, $n = 1, 2, \ldots, N$, then we expect

$$\frac{A_1 + A_2 + \ldots + A_N}{N} \to \langle \hat{A} \rangle \qquad \text{for } N \to \infty \tag{5.133}$$

with probability one. Then $\langle \hat{A} \rangle$ is indeed the expectation value of the observable $\hat{A}$ if we interpret the measurement as a statistical experiment. The term "expectation value" is then justified. An expectation value can always be written as

$$\langle \hat{A} \rangle = \sum_A p_A \, A, \tag{5.134}$$

where p_A with $\sum_A p_A = 1$ is the probability of measuring the value A. Comparing this to Eq. (5.132) shows that

$$p_A = |\langle \psi_A, \psi \rangle|^2. \tag{5.135}$$

This relation is called the *Born rule*. This rule cannot be derived from the Schrödinger equation, nor does it follow from von Neumann's projection postulate. The latter only tells us which measured values are possible and in which state our system will be after the measurement. The Born rule is thus an additional postulate of mordern quantum theory. Conceptually speaking, this postulate states that the probability interpretation is correct and how one can calculate the probabilities for the results of measurements.

Can the Born rule be confirmed experimentally? Surely not by a single measurement since it is a statistical statement. But together with the standard frequentist interpretation of probabilities, the Born rule makes a prediction for a series of many measurements: Assuming we take $N \gg 1$ measurements of $\hat{A}$ with the same conditions and N_A times measure the value A. Then we expect that the fraction N_A/N converges towards p_A for large N,

$$\lim_{N \to \infty} \frac{N_A}{N} = p_A. \tag{5.136}$$

For the debate on the interpretation of quantum mechanics, tt is important to mention that these statements are only valid "with probability one". It *could* also happen that in such a series of experiments N_A/N converges towards a different value, but the larger the N, the less often will this occur.

The Born rule is very well confirmed in the described sense. There is an experiment that works with the naked eye and without any special equipment: According to Born, four candles bundled closely together in a distance of 200 m to the observer are as bright as a single candle in a distance of 100 m. In more detail, we can interpret the classical electromagnetic field as the wave function of photons (to be exact, it is the four-vector potential (A^μ)), as one can see in field theory. The analogue to the Schrödinger equation is the wave equation for the field. It predicts that the amplitude of the field of a point source decays as $1/r$ with the distance r. From a classical viewpoint, it must be so because the intensity decays with $1/r^2$. The surface of an imagined sphere with a radius of r around the point source is of course proportional to r^2. Hence, the energy current through the sphere's surface is independent of its radius, which must be the case due to conservation of energy.

We now turn to the quantum-theoretical perspective. Seeing is actually a position measurement of photons. We sense light if a photon is absorbed by a receptor molecule in the retina of our eye. This happens with a certain probability p. According to Born, this probability is proportional to the absolute value squared of the field and thus, for a point source, to $1/r^2$. If we were to observe a different dependency (including the possibility that p doesn't depend on the distance), then the Born rule would be experimentally falsified. The Born rule predicts that the probability of sensing photons from a point source decays with a factor of $1/4$ if the source is at twice the distance. The essential idea would then be that multiple, say n times, the same type of light source (candles or multiple LEDs of the same kind) are so close together that the eye lacks the resolution to see each source individually, which makes it seem like a single light source that emits n times the number of photons. If we observe one point source in the distance l and compare it to four point sources in a distance of $2l$, then, according to Born, we should be able to perceive the same amount of photons from both arrangements on average over a sufficiently long time ($N \gg 1$). Basically, they will seem equally bright. Since a lot of photons are received it doesn't take long to see this.

5.7 The Time-Independent Schrödinger Equation

We will limit ourselves to cases where the Hamilton operator does not explicitly depend on time. Then we can rewrite the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}\, \psi(\vec{r}, t) = \hat{H}\, \psi(\vec{r}, t) \tag{5.137}$$

in a simpler form using a *product ansatz* or *separation ansatz*. We thus set

$$\psi(\vec{r}, t) = \psi(\vec{r})\, f(t). \tag{5.138}$$

It then follows that

$$i\hbar\, \psi(\vec{r})\, \frac{df}{dt} = \left[\hat{H}\psi(\vec{r})\right] f(t). \tag{5.139}$$

We divide by $\psi(\vec{r}, t) = \psi(\vec{r})\, f(t)$ and obtain

$$i\hbar\, \frac{\dot{f}(t)}{f(t)} = \frac{\hat{H}\psi(\vec{r})}{\psi(\vec{r})}. \tag{5.140}$$

We should check afterwards what happens at points with $\psi(\vec{r}, t) = 0$, where we can't divide as we do here. Now the left-hand side is independent of $\vec{r}$ and the right-hand side is independent of t but both sides are equal for all $\vec{r}$ and t. Hence, the same value is independent of $\vec{r}$ and of t and is thus a constant that we call a *separation constant*. Here, this constant is denoted by E. It then follows that

$$i\hbar\, \frac{\dot{f}(t)}{f(t)} = E, \tag{5.141}$$

$$\frac{\hat{H}\psi(\vec{r})}{\psi(\vec{r})} = E \tag{5.142}$$

and thus

$$i\hbar\, \frac{df}{dt} = E f(t), \tag{5.143}$$

$$\hat{H}\psi(\vec{r}) = E\psi(\vec{r}). \tag{5.144}$$

Evidently, E is of the dimension of energy. Equation (5.143) has a single linearly independent particular solution $e^{-iEt/\hbar}$ and the corresponding general solution is

$$f(t) = f_0\, e^{-iEt/\hbar}. \tag{5.145}$$

This is true for any complex number E. Of course, for physically meaningful solutions E must be real since otherwise $f(t)$ would contain an exponentially growing or decaying factor, which would violate the normalization condition $\int d^3r\, |\psi|^2 = 1$, which must hold at all times t. Apart from this, Eq. (5.143) doesn't limit possible values of E. Equation (5.144) is the *time-independent Schrödinger equation*,

$$\hat{H}\, \psi_n(\vec{r}) = E_n\, \psi_n(\vec{r}). \tag{5.146}$$

This is an *eigenvalue equation* for the eigenvalues (in this case *eigenenergies*) E_n and the corresponding eigenfunctions $\psi_n(\vec{r})$ of the Hamilton operator $\hat{H}$. Since the energy must be real—firstly because it is a measurable quantity and secondly to ensure that the solution of the time-dependent Schrödinger equation stays normalizable—it is sensible for the Hamilton operator $\hat{H}$ to be hermitian. (Non-hermitian Hamiltonians are occasionally considered as well, e.g., for open systems.) It can happen that certain eigenvalues are *degenerate*, meaning that one value E appears multiple times in he series $E_1, E_2, \ldots$ As discussed in the last section, one can construct an orthonormal basis from the eigenfunctions of a hermitian operator. Because of this, we can expand every square-integrable function in terms of these eigenfunctions. This is especially useful for the solution of the *time-dependent* Schrödinger equation: The particular solutions of this equation can be constructed from the solutions of the separated equations for $\psi(\vec{r})$ and $f(t)$. The resulting particular solutions are

$$\psi_n(\vec{r},t) = \psi_n(\vec{r})\, e^{-iE_n t/\hbar}. \tag{5.147}$$

The general solution is found by superposition of particular solutions as

$$\psi(\vec{r},t) = \sum_n a_n\, \psi_n(\vec{r})\, e^{-iE_n t/\hbar}, \tag{5.148}$$

with arbitrary coefficients $a_n \in \mathbb{C}$ (the condition that $\psi(\vec{r},t)$ is normalized leads to a condition for the a_n).

A typical problem is to find $\psi(\vec{r},t)$ for $t > 0$ for a given system, described by a Hamilton operator $\hat{H}$, if $\psi(\vec{r},0)$ is given as an initial condition. This means that the system is prepared in a certain state at a certain time, described by the wave function, and we are interested in its evolution for later times. To solve this initial-value problem, we decompose $\psi(\vec{r},0)$ into eigenfunctions of $\hat{H}$,

$$\psi(\vec{r},0) = \sum_n a_n\, \psi_n(\vec{r}). \tag{5.149}$$

Using the orthonormality of the system, we can find the coefficients a_n of the eigenfunctions,

$$\int d^3r\, \psi_n^*(\vec{r})\, \psi(\vec{r},0) = \sum_{n'} a_{n'} \int d^3r\, \psi_n^*(\vec{r})\, \psi_{n'}(\vec{r}) = \sum_{n'} a_{n'} \delta_{nn'} = a_n, \tag{5.150}$$

so that

$$a_n = \int d^3r\, \psi_n^*(\vec{r})\, \psi(\vec{r},0) = \langle \psi_n, \psi(t=0) \rangle. \tag{5.151}$$

The solution for any time is then

$$\psi(\vec{r},t) = \sum_n \langle \psi_n, \psi(t=0) \rangle\, \psi_n(\vec{r})\, e^{-iE_n t/\hbar}. \tag{5.152}$$

Keep in mind that this relatively easy way of solving the problem only works since the Schrödinger equation is linear.

Now, as an example, we consider a one-dimensional box of length L and with impregnable walls at $x = 0$ and $x = L$. It is plausible and will be justified later that $\psi(x,t)$ must disappear at the edges. The potential $V(\vec{r})$ disappears in this case inside the box and the Hamilton operator takes the form

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}. \tag{5.153}$$

As an initial condition, we consider the wave function

$$\psi(x,0) = c\, \sin^3 \frac{\pi x}{L} \tag{5.154}$$

at the time $t = 0$, where c is a normalization constant (one can calculate that $c = 4/\sqrt{5L}$). The function $\psi(x,0)$ evidently fulfills the boundary conditions at $x = 0$ and $x = L$.

Step 1: Eigenfunctions $\psi_n(\vec{r})$ and eigenenergies E_n. Die time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2}\, \psi_n(x) = E_n \psi_n(x), \tag{5.155}$$

with the boundary conditions $\psi_n(0) = \psi_n(L) = 0$. We can also write the equation as

$$\psi_n''(x) = -\frac{2mE_n}{\hbar^2}\, \psi_n(x). \tag{5.156}$$

Ansatz:

$$\begin{aligned}
& \psi_n(x) = Ae^{ikx} + Be^{-ikx} && (5.157)\\
\Rightarrow\quad & \psi_n''(x) = -Ak^2e^{ikx} - Bk^2e^{-ikx} = -k^2\psi_n(x) && (5.158)\\
\Rightarrow\quad & k^2 = \frac{2mE_n}{\hbar^2} && (5.159)\\
\Rightarrow\quad & E_n = \frac{\hbar^2k^2}{2m}. && (5.160)
\end{aligned}$$

Boundary conditions:

$$\psi_n(0) = A + B = 0 \tag{5.161}$$

and

$$\begin{aligned}
& \psi_n(L) = Ae^{ikL} + Be^{-ikL} = 0 && (5.162)\\
\Rightarrow\quad & B = -A \quad \text{und} \quad \psi_n(L) = A(e^{ikL} - e^{-ikL}) = 2iA\sin kL = 0 && (5.163)\\
\Rightarrow\quad & kL = n\pi, \quad n \in \mathbb{N}. && (5.164)
\end{aligned}$$

Thus the eigenenergies are

$$E_n = \frac{\pi^2\hbar^2}{2mL^2}\,n^2, \quad n \in \mathbb{N} \tag{5.165}$$

and the corresponding eigenfunctions are

$$\psi_n(x) = c_n \sin\frac{n\pi x}{L}, \tag{5.166}$$

where c_n are normalization constants. We find that

$$\begin{aligned}
& \int_0^L dx\,|\psi_n(x)|^2 = |c_n|^2 \int_0^L dx\,\sin^2\frac{n\pi x}{L} = |c_n|^2\,\frac{L}{2} \overset{!}{=} 1 && (5.167)\\
\Rightarrow\quad & |c_n|^2 = \frac{2}{L}. && (5.168)
\end{aligned}$$

We choose $c_n = \sqrt{2/L}$ so that

$$\psi_n(x) = \sqrt{\frac{2}{L}}\,\sin\frac{n\pi x}{L}. \tag{5.169}$$

Step 2: Expansion of the initial condition. We have

$$\begin{aligned}
a_n &= \int_0^L dx\,\psi_n^*(x)\,\psi(x,0)\\
&= \int_0^L dx\,\sqrt{\frac{2}{L}}\,\sin\left(\frac{n\pi x}{L}\right)\,\frac{4}{\sqrt{5L}}\,\sin^3\frac{\pi x}{L}\\
&\overset{\text{table}}{=} \frac{4\sqrt{2}}{\sqrt{5}L} \times \begin{cases} \frac{3L}{8} & \text{for } n = 1\\ -\frac{L}{8} & \text{for } n = 3\\ 0 & \text{otherwise}\end{cases}\\
&= \begin{cases} \frac{3}{\sqrt{10}} & \text{for } n = 1\\ -\frac{1}{\sqrt{10}} & \text{for } n = 3\\ 0 & \text{otherwise.}\end{cases} && (5.170)
\end{aligned}$$

(The integral can also easily be solved by expressing the sinus functions as complex exponential functions.) We obtain

$$\psi(x,0) = \frac{3}{\sqrt{10}}\,\psi_1(x) - \frac{1}{\sqrt{10}}\,\psi_3(x). \tag{5.171}$$

Step 3: Put together the solution. For all t, it holds that

$$\begin{aligned}\psi(x,t) &= \frac{3}{\sqrt{10}}\,\psi_1(x)\exp\left(-i\,\frac{\pi^2\hbar}{2mL^2}\,t\right) - \frac{1}{\sqrt{10}}\,\psi_3(x)\exp\left(-i\,\frac{9\pi^2\hbar}{2mL^2}\,t\right)\\ &= \frac{3}{\sqrt{5L}}\,\sin\frac{\pi x}{L}\,\exp\left(-i\,\frac{\pi^2\hbar}{2mL^2}\,t\right) - \frac{1}{\sqrt{5L}}\,\sin\frac{3\pi x}{L}\,\exp\left(-i\,\frac{9\pi^2\hbar}{2mL^2}\,t\right).\end{aligned} \tag{5.172}$$

Here we show a plot of $|\psi(x,t)|^2$ for $L = 1$, where the time t progresses from 0 (blue) to $m/4\pi\hbar = m/2h$ (red):