Lectures on quantum mechanics for material scientists

Educational Material

Author(s):
Pescia, Danilo

Publication date:
2023

Permanent link:
https://doi.org/10.3929/ethz-b-000621567

LECTURES ON QUANTUM MECHANICS FOR MATERIAL SCIENTISTS

by

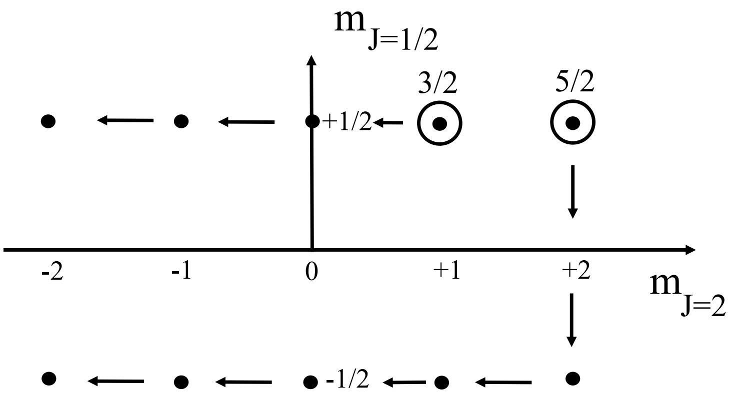
D. Pescia

Laboratory for Solid State Physics, Department of Physics

ETH Zurich

CH-8093 Zurich

pescia@phys.ethz.ch



Preface

When speaking of quantum mechanics, one has typically a physics or chemistry student in mind. However, research on materials is now so advanced that students of material science must be trained to understand and above all use those deepest concepts of quantum mechanics that were reserved to physics students in the past. It is along these lines that these lectures discuss the fundamental concepts of modern Quantum Physics such as the failure of classical mechanics, the postulates and mathematical structure of Quantum Physics, the electronic structure of atoms, molecules and solids, the electron spin and time dependent perturbations. The lectures are constructed in such a way that the various concepts are introduced by starting from the explicit and detailed treatment of a simple example. These particular cases are then generalized and put into a ”moderately ” rigorous framework – the word ”moderately” meaning that the mathematics is not worked out in details but the essential steps are rigorous enough to withstand practical applications. The present manuscript summarizes the content of the various lectures. The lectures were distributed over a semester at ETH Zurich (about 13 weeks, four hours per week). The course included a set of problems that were intended to complement the lectures themselves with concrete examples and that were solved weekly by the students (one hour per week). I am grateful to the assistants (in particular, for the final version, A.-K. Thamm and M. Demydenko) for mentoring the students during the exercises. I am also grateful to Pascal Studer for providing an excellent graphical version of the various figures. The manuscript contains a larger amount of information than the lectures: in my opinion, this provides the readers with the opportunity to finding out by themselves the essential aspects and, ultimately, to acquire a real understanding of the matter. Those section marked with * are for advanced reading. Some of the experiments presented contain a link to the site ”Vorlesungsexperimente Departement Physik” of ETH Zurich. The team that prepared these experimental demonstrations over the years has supported the present lectures with a large amount of experiments that could be performed in-situ and in real time. They also made their work available online. I am grateful to the team for their work in supporting these lectures in particular and the entire physics lecture businnes at ETH Zurich.

The number of textbooks on Quantum Mechanics is, of course, approaching ”infinity”. One may ask: why such a further manuscript. The best answer is: this manuscript is unnecessary. A less straighforward answer is: when a manuscript is based on actual lectures, it takes into account the striving of the teacher to present the various concepts in such a way that they are understandable and can be put into practice. Material scientists need exactly this kind of approach. I hope that this manuscript satisfies their needs for having a piece of practical quantum mechanics they can refer to in their further career.

The knowledge used to prepare these lectures was acquired by reading a large amount of literature on the subject. In particular, I would like to quote the lectures on quantum mechanics held by the late Prof. R. Jost at ETH Zurich during the academic year 1977-1978, at which I had the honor to participate as a student. Furthermore, I found the manuscript by the late Prof. W. Hunziker inspiring. The present lectures show my personal understanding of the subject. This understanding also emerged during the interaction with the many students that frequented my courses. A thank goes also to them.

D. Pescia, Department of Physics, ETH Zurich, March 2023

Contents

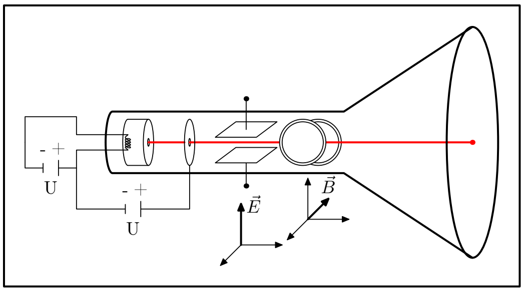
Part I.

Schrödinger Quantum Mechanics

1. The origin of quantum mechanics

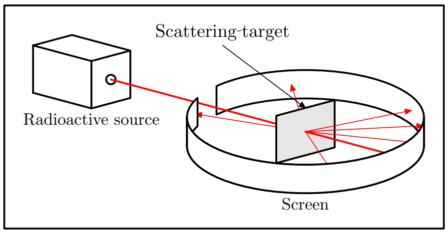
1.1. The year 1911: J.J. Thomson and E. Rutherford

At the beginning of the twentieth century, numerous observations made it clear that Newtonian mechanics was not able to explain important experimental results. The most important discoveries concern the concept and the experimental evidence about ”Atoms”. Democritus, some 24 centuries ago, (460-371 BC) interpreted, correctly, the transformation of water from liquid to vapour as evidence that matter consists of small, indivisible units that he called ”atoms”. This awareness was more philosophical and only survived in the shadow of the doctrin of Aristoteles, according to which matter was a combination of the four elements that, in the Middle Age, formed the basis of alchemy: water, earth, fire, air. Exerimental evidence for the existence of atoms came from the chemists, after Pierre Gassendi in 1647 and Robert Boyle in 1661 developed further the theory of Democrit, arguing that matter is composed of combinations of atoms each carrying a different shape, dimension and weight, i.e. each forming different elements – but certainly not the four elements of alchemy. Antoine Lavoisier, considered the founder of modern chemistry, introduced, in 1789, the current concept of the chemical element and identified some of them (oxygen, hydrogen, sulfur (soufre), phosphorus, carbon (carbone) antimony, silver etc.). It was John Dalton, in 1803, that introduced atoms definitely in chemistry as the only possible explanation of the fact that elements react always in quantitative proportions of some integers to form compounds: each element consists of similar atoms, which then undergo chemical reactions to build compounds. The contribution of physics followed much later: J.J. Thomson observed in 1897 the motion of so-called ”cathode rays” in the presence of electric and magnetic fields and found that they behave like negatively charged particles.



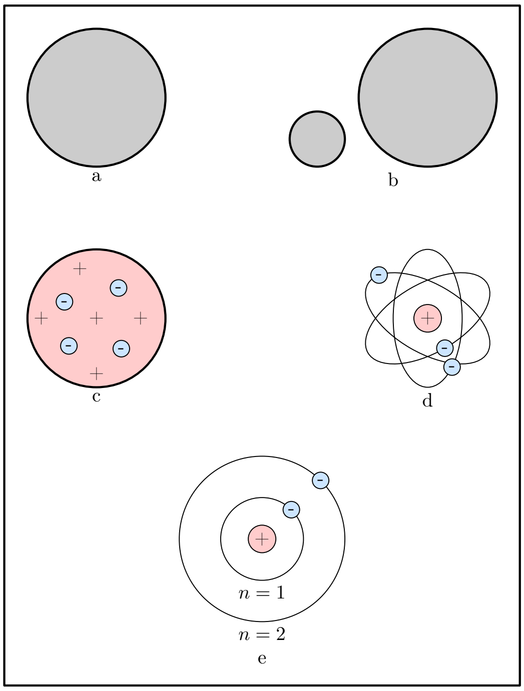
Experiment for determining $\frac{e}{m}$*, carried out by J.J. Thomson in 1897. The electrons, generated by thermo-emission from a hot wire (the cathode, left), are accelerated by a voltage* U *between cathode and anode. When traversing a plate capacitor, they are deflected along the vertical direction by an electric field* $\vec{E}$*. If a coil applies a magnetic field* $\vec{B}$ *along the horizontal direction, the deflection caused by* $\vec{E}$ *can be compensated so that the electrons (red beam) pass the arrangement exactly parallel along the axis of the instrument (perpendicular to both* $\vec{E}$ *and* $\vec{B}$*).* $\frac{e}{m}$ *can be derived from the classical relation* $\frac{1}{2U}\frac{E}{B}^{2}$.

Thomson determined $\frac{e}{m}$ for the electron and recognized it as a particle. From this he constructed the first model of the atom, consisting of small point-like particles (the electrons) embedded in a positively charged, possibly spherical, ’sponge’ (Thomson model). Later, in 1909, E. Rutherford scattered α-particles at a very thin gold foil and found a sizeable number of particles being backscattered.



α-particles are positively charged and are produced by radioactive decay. In the Rutherford scattering experiment, a beam of α particles is scattered onto a metal foil (typically gold). The α particles scattered by the ”target” are registered on a screen. Rutherford found that most particles pass undisturbed through the foil, but some are scattered at great angles (even backwards). This speaks against the Bohr model, because the scattering potential of a positively charged sphere with the extension of an atom would hardly have any great effect on the particles and would certainly not have scattered them backwards. In contrast, a model of the atom as an empty space with a positively charged nucleous in the middle is more compatible with the experimental results of Rutherford.

This sizable number of backscattered α particles was puzzling. Suppose that the positive atomic charge would extend over the entire size of atoms – estimated to 0.1 nm. Then the deflection of the α-particles by Coulomb repulsion acting at such distances was estimated by Rutherford to be almost negligible. A much smaller size of the positive charge - the emergence of a ”nucleous” with a radius less that 10^{-5} nm – was necessary to explain the backscattered particles, indicating that most of the mass of the atom is concentrated in a very small charged atomic nucleous in the center of the atom (Rutherford’s atomic model), around which the electrons under the influence of the Coulomb force circle as point particles. The Rutherford model had a drawback: orbiting electrons around a nucleous fall into the nucleous by emission of radiation after a very short time. This was the end of the line for classical models of the atom that included Coulomb interaction, Newton mechanics (leading to centrifugal forces that prevented electrons to fall into the nucleous) and classical electrodynamics.



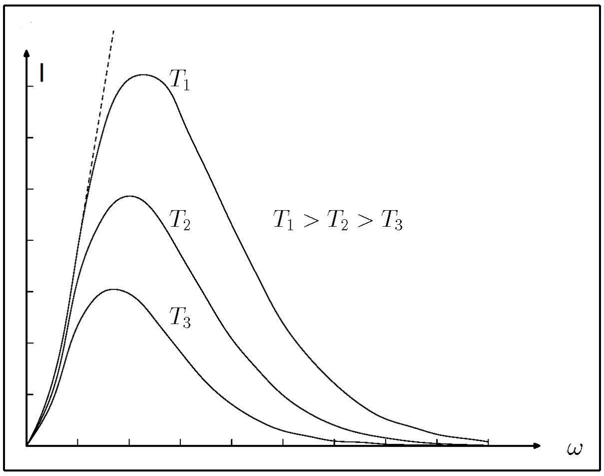
A summary of the models of atoms. Democritus (a) considered matter as consisting of small, indivisible particles. Dalton (b) suggested that matter consists of indivisible atoms, the atoms of different elements always differ in their mass and size. Moreover, he proposed that an integer number of atoms of different elements form compounds. Thomson (c) recognized that the electrons are components of the atoms and can escape from matter ("cathode rays"). Each atom consists of an electrically positively charged sphere, in which electrically negatively charged electrons are stored - like raisins in a cake. Rutherford (d) summarized the findings from his experiment as an atom model: in Rutherford's atom, the positive charge is concentrated within a point (the nucleous) that is in the atomic center. The diameter of the atomic nucleous is only one ten-thousandth of the total atomic diameter. The atomic shell contains the negatively charged electrons that revolve around the atomic nucleous. The atomic shell is almost empty because the electrons are much smaller than the atom. Bohr (e) (see later) proposed circular orbits for the electrons that develop with a discrete ("quantized") set of angular momenta. Electrons are "protected" by the quantization to decay into the nucleous.

1.2. The year 1900: Planck.

Any body at finite temperature emits radiation in a continuum range of frequencies. An atomic model of the source of this radiation foresees that the atoms building the body can be modelled as an ensemble of harmonic oscillators (possibly interacting, but this is not an essential element of the model). We know now that the frequency of the oscillations can vary over a certain continuous range (depending on the material properties of the medium), but this is also a non-essential element of the model. What is essential is the fact that the oscillators at a finite temperature can assume different total amplitudes of oscillation – the probability that a certain amplitude A is realized being proportional to

$$e^{-\frac{\frac{1}{2}\cdot m\cdot\Omega^2\cdot A^2}{k_B\cdot T}}$$

In this expression we use, for the sake of simplicity and to stick with the conventional wisdom prior to the year 1900, $\frac{1}{2}\cdot m\cdot\Omega^2\cdot A^2$ to express the energy of the harmonic oscillator – this is therefore the classical energy of a single oscillator. m is the mass of the oscillating particle and Ω a characteristic frequency of oscillation. A^2 can vary in a continuous manner. T is the equilibrium temperature and k_B is the Boltzmann constant, that translates temperatures into energies. The exponential law for the probability is an exact result of equilibrium thermodynamics of macroscopic bodies and is still a valid one – its is called the Boltzmann-Gibbs or canonical distribution. To understand the origin of radiation, is enough to think that the equilibrium state is one where the individual oscillators might assume amplitudes which differ for short times from the thermal average $<A^2>$ but decay, after enough long times, to $<A^2>$. The body therefore is seen as a macroscopic set of jittering oscillators which might absorb or emit radiation is the process of keeping the equilibrium value $<A^2>$. By this absorpion and emission of electromagnetic waves the body becomes a container of a further system besides a gas of harmonic oscillators: a system consisting of electromagnetic radiation with a certain energy density. This model was of course very clear to all involved in measuring and computing the radiation emitted by a body at finite temperature. Rayleigh and Jeans, for instance (L. Rayleigh: Remarks upon the Law of Complete Radiation, Phil. Mag. Series 1, **49**, 1900, p. 539–540; J. H. Jeans, On the partition of energy between matter and Aether, Phil. Mag. Series 6, **10**, 1905, p. 91–98) computed the spectral energy density of the radiation field as a function of the radiation frequency ω using a classical model in which the amplitude A^2 is taking continuous values and found that the intensity of the radiation at a given temperature was increasing as ω^2 (we will find later this result, using a model of the oscillators proposed by Einstein). This represented a clear contrast with experiments at high frequencies. These experiments measured the intensity of the emitted "black body" radiation (this intensity being related to the energy density of the electromagnetic radiation in the body) and found that the initial ω^2-dependence turned into a decreasing behaviour at higher frequencies.



The emitted radiation as a function of the radiation frequency is expected in the classical approach to grow as ω^2 *(dashed line). Experimentally, the graph follows closely the functional dependence on* ω *computed by Planck (solid lines), shown for different temperatures* T*.*

It was Planck (M. Planck, Faksimile aus den Verhandlungen der deutschen Physikalischen Gesellschaft, **2**, (1900) p.237-245, Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum) who found an explanation for this deviation from the classical ω^2-dependence: He assumed that oscillations could only take place with some discrete amplitudes. More precisely: he assumed that the energy of oscillators could only assume discrete values E_n with $E_n = \text{const} \cdot n \cdot \Omega$, with $n = 0, 1, 2, \ldots$. He then computed the frequency dependence of the energy density of the radiation field (we will obtain Planck result later, using a method proposed by Einstein) in thermal equilibrium, and adjusted the proportionality constant arising in $E_n = \text{const} \cdot n \cdot \Omega$ and appearing in the equilibrium energy density in order to **fit** the shape of the experimentally measured ω-dependence. He found that the proportionality constant assumes the value which is now known as Planck constant of action

$$\begin{aligned} \hbar &= 1.04 \cdot 10^{-34} \text{Joule} \cdot \text{seconds} = 6.582 \cdot 10^{-16} \text{eV} \cdot \text{seconds} \\ h &\doteq 2\pi \cdot \hbar = 6.626 \cdot 10^{-34} \text{Joule} \cdot \text{seconds} = 4.136 \cdot 10^{-15} \text{eV} \cdot \text{seconds} \end{aligned}$$

1.3. The year 1905: Einstein.

A system of electromagnetic radiation at a given frequency ω and given temperature T has not only an energy density, given by Planck formula, but, as recognized by Einstein, it must be assigned, from a thermodynamic point of view, an entropy as well. Einstein computed the entropy of the radiation in a frequency interval $[\omega, \omega + d\omega]$ as amounting to

$$k_B \cdot \frac{U}{\hbar\omega} \cdot \ln(\frac{V}{V_0})$$

with U being the total energy of the radiation in the interval $[\omega, \omega + d\omega]$ and V the volume of the system. V_0 is some constant that, for the purpose of finding the significance of this formula, does not play a role. This formula is very much telling when compared to the entropy of an ideal gas of N independent particles:

$$k_B \cdot N \cdot \ln(\frac{V}{V_0})$$

Einstein concluded, by comparing the two formulas, that monochromatic light behaves from the thermodynamical point of view just as if it were made up of mutually independent ”particles” each carrying the energy

$$E = \hbar \cdot \omega$$

In 1917 Einstein proposed that the ”particles” be assigned a ”momentum” as well, to be computed from the wavenumber k of the light (recall that $k = \frac{2\pi}{\lambda}$, λ being the wavelength of the radiation) acccording to

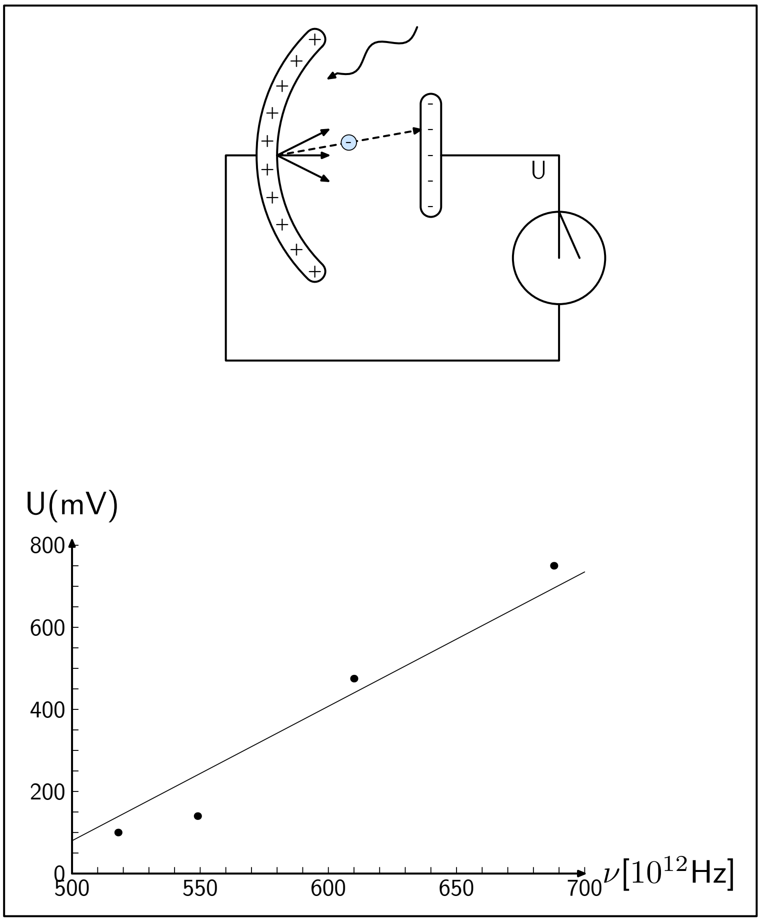
$$p = \hbar \cdot k$$

Nowadays we call these ”Einstein particles” ”photons” and Einstein hypothesis is called the ”photon hypothesis’.

In his work, Einstein proceeds to find an application of this hypothesis: he proposes that the photon hypothesis explains the (so far) mysterious photoelectric effect, discoverd by Hertz in 1887 and quantitatively underlying to experiments by Lenard in 1902. Lenard found that the maximum kinetic energy of electrons emitted by illuminating a metal with light does not depend on the intensity of the radiation but on the frequency of the light. Einstein proposed to use the photon hypothesis to explain this behaviour and found that the experimental data were properly described by the particle like energy conservation relation

$$E_{max} = \hbar\omega - \varphi$$

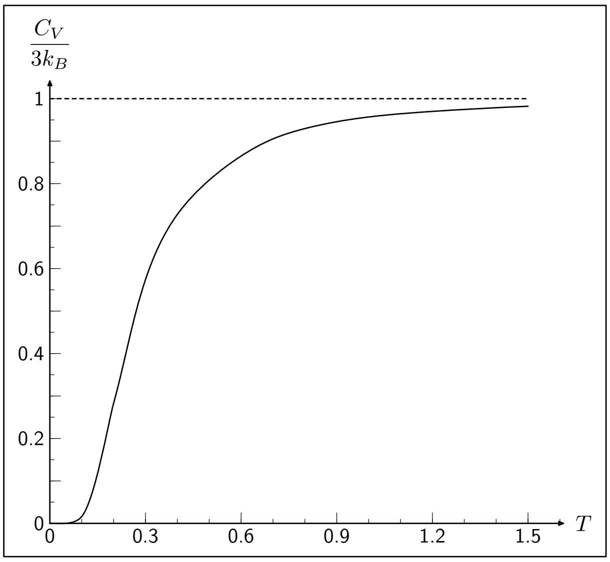
e.g. the slope of the graph E_{max} versus ω should be just $\hbar$. The constant φ is a minimum energy barrier that the electrons have to overcome when exciting the plate (the so-called work function) and is typically 4.5 eV.



The photoelectric effect is illustrated in the Figure. A simple experimental arrangement which can be used to measure the photoeffect consists of a metallic plate (the cathode) and a collector (the anode). If light hits the cathode at a certain frequency, electrons leave the cathode at suitable light frequencies and are collected by the collector, which charges negatively. This creates an electric field that slows down the electrons. The electric field continues to grow until the potential energy difference U between the cathode and the anode equals the maximum kinetic energy of the electrons. If we plot the voltage between the cathode U and the anode as a function of the light frequency, we find that the data lie on a straight line $U = \nu \cdot \frac{2\pi\cdot\hbar}{e} - \frac{\varphi}{e}$*. From the slope of the line,* $\hbar$ *can be determined by known electron charge. If one try to excite electrons with a frequency below that of the photothreshold* $\frac{\varphi}{\hbar}$*, no photoemission current is recorded, even if the intensity of the light is increased. See* `https://experimente.phys.ethz.ch/de/100/10002/20071/30457/`

1.4. The year 1906: Einstein.

The ensemble of harmonic oscillators co-exsiting with its radiation field in a medium has itself an equilibrium total energy $E(T,V)$ and a specific heat (or heat capacity) $C_V = \frac{\partial E}{\partial T}\,|_V$ which can be recorded as a function of temperature by – at that time – standard calorimetry. It was one of the most celebrated results of thermodynamics applied to a gas of classical oscillators that the total equilibrium energy per oscillator, computed assuming the classical mechanical energy $\frac{1}{2} \cdot m \cdot \Omega^2 \cdot A^2$ – originating from Newton Mechanics – is just $3 \cdot k_B \cdot T$, the number 3 arising from the fact that an oscillator in a three dimensional body has three degrees of freedom. This is the result of a strict theorem of classical statistical mechanics known as the equipartition theorem, which assigns to each degree of freedom subject to potentials energy a total energy of $k_B \cdot T$. This total energy produces a Dulong-Petit contribution $3 \cdot k_B$ to the specific heat in solids which is, most remarkably, temperature independent and universal. The experimental results, instead, show the completely different picture of a temperature dependent, non-universal heat capacity in common metals. The explanation for the experimental behaviour residing outside the equipartition theorem was provided by A. Einstein, assuming – along the lines of reasoning by Planck – that the solid contains a gas of harmonic oscillators with a characteristic frequency Ω and allowed to assume only discrete values of energies, given by $n \cdot \hbar \cdot \Omega$, $n = 0, 1, 2, 3, ...$



Sketch of the the temperature dependence of the specific heat as measured in various solids. The temperature dependence varies from material to material, but, if one divides the real temperature by a a material specific characteristic temperature, the experimental data from different materials fall approximately onto the same curve, sketched in the figure. The theoretical curve obtained here below in this section fits well the experimental data. The dashed horizonal line is the purely classical Dulong-Petit result.

* **Einstein model of the specific heat in solids.** It is educational to follow in details the computation of the total energy along the two lines of reasoning – the classical one, leading to the equipartition – and the one assuming quantized energy levels, leading to a pronounced deviation from classical equipartition. For this purpose we summarize some fundamental results of statistical physics. At finite temperarture, a system in thermodynamic equilibrium contains a set of excitations which are typically characterized by some excitation energy E_n. The probability that such an excited state is realized is given by the Boltzmann-Gibbs distribution

$$P_n = \Omega_n \cdot \frac{e^{-\frac{E_n}{k_B \cdot T}}}{\sum_n \cdot \Omega_n \cdot e^{-E_n/k_B T}}$$

Ω_n is the degeneracy of the excited state with energy E_n. For instance, a given energy state $\frac{1}{2}\cdot m\cdot\Omega^2\cdot A^2$ is at least twice degenerate, taking into account that a state with amplitude $+A$ and $-A$ have the same excitation energy. The denominator is the sum over all possible excited states and normalizes the probability of assuming any excited state to 1. Once P_n is known, one can compute thermal averages, such as the average energy at a temperature T by the recipe

$$<E>|_{T,V,N}=\frac{\sum_n E_n\cdot\Omega_n\cdot e^{-E_n/k_BT}}{\sum_n\Omega_n\cdot e^{-E_n/k_BT}}$$

These rules govern equilibrium statistical mechanics. The problem of computing thermal averages reduces to compute the energies of all possible excited states – a problem which is, in general, extremely difficult to solve with the exception of very few special models. The best one can do is to find a set of relevant excitations and hope that the thermal average computed over them is representative. Notice that we have chosen a formulation with discrete energy levels, which is strictly speaking only correct in quantum mechanics. When Newton mechanics governs the motion, one has a continuum set of excitation energies and the formulation has to be modified accordingly. The lore of classical statistical mechanics foresees that the possible states of a classical system are described with the phase space of the variables q - the coordinates of the space where the system can reside (the so-called configuration space) - and p - the possible values of the momenta that the system can assume. The total energy is described by the Hamiltonian function $H(q,p)$ which contains e.g. the kinetic energy and potential energies of the various interactions. The q,p parameter space is a continuum one so that the Boltzman-Gibbs probability distribution translates into an integral:

$$\sum_n(\ldots)\longrightarrow\int(\ldots)\rho(q,p)dq\cdot dp$$

The number of possible states in the intervall $(q+dq,p+dp)$ is the density of states $\rho(q,p)$. This quantity is ill- defined in classical statistical physics but in order to recover results of statistical physics based on classical mechanics starting from results based on quantum mechanics one can set $\rho(q,p)=\frac{1}{h^d}$, d being the dimensionality of the system. This means that the phase space is thought of an ensemble of cells with volume h^d and containing 1 possible state of the system. The probability of finding the system in the interval state $(q+dq,p+dp)$ does not contain $\rho(q,p)$ explicitely and writes

$$P(q+dq,p+dp)=\frac{e^{-\frac{H(q,p)}{k_B\cdot T}}\cdot dq\cdot dp}{\int dq\cdot dp\cdot e^{-H(q,p)/k_BT}}$$

We would like to use this expression to compute the average energy $e(T)$ of one single classical harmonic oscillator with one degree of freedom, i.e. with $H(q,p)=\frac{p^2}{2m}+\frac{m\Omega^2q^2}{2}$. Of course, a real system hosts many of such oscillators, but if they are not interacting one can deal with one single particle and the thermal averages for the real system are obtained by multiplying the thermal average of one particle with the number of particles. We find that Gaussian integrals appear in the computation, which gives finally the classical equipartition result:

$$e(T)=\int_{-\infty}^{\infty}dq\int_{-\infty}^{\infty}dp\cdot(\frac{p^2}{2m}+\frac{m\Omega^2q^2}{2})\cdot P(q+dq,p+dp)=k_B\cdot T$$

with

$$P(q+dq,p+dp)=\frac{e^{-\frac{\frac{p^2}{2m}+\frac{m\Omega^2q^2}{2}}{k_B\cdot T}}\cdot dq\cdot dp}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}dq\cdot dp\cdot e^{-(\frac{p^2}{2m}+\frac{m\Omega^2q^2}{2})/k_BT}}$$

We now assume that the energies of the harmonic oscillator are quantized to

$$\epsilon_n=n\cdot\hbar\cdot\Omega$$

Later, when dealing with the quantum mechanical harmonic oscillator, we will prove this assumption starting from "first principles". $e(T)$ writes now

$$e(T) = \frac{\sum_{n=0}^{\infty} n\cdot\hbar\Omega\cdot e^{-\hbar\Omega n/kT}}{\sum_{n=0}^{\infty} e^{-\hbar\Omega n/kT}}$$

where the sum runs over the discrete index n. To evaluate the sums, we write

$$\hbar\Omega\frac{\sum_n n\cdot x^n}{\sum_n x^n} \qquad x \doteq e^{-\hbar\Omega}$$

and use the identities

$$\sum_n x^n = \frac{1}{1-x} \qquad \sum_n n\cdot x^n = x\frac{d}{dx}\sum_n x^n = \frac{x}{(1-x)^2}$$

so that we obtain

$$< e(T) >= \hbar\Omega\cdot\frac{1}{e^{\hbar\Omega/kT}-1}$$

and

$$c_V(T) = k_B\frac{\frac{(\hbar\Omega)^2}{k_B\cdot T}\cdot e^{\frac{(\hbar\Omega)^2}{k_B\cdot T}}}{e^{\hbar\Omega/k_BT}-1}$$

This result, obtained first by Einstein, shows two important facts: First, the specific heat contains the material parameter Ω, which makes it material- dependent. Second, the temperature dependence of the specific heat, sketched in the previous figure, is not just a straight horizontal line and is found to better fit the experimental data.

The understanding of the experimental specific heat data was an important achievement that supported the original intuition by Planck of quantized energies for harmonic oscillators. One important comment: for large temperatures, $e(T)$ approaches the classical limit $k_B\cdot T$. We point out that the cross-over to classical behavior at high temperatures is a generally valid trend in matter and suggests that quantum behavior in matter is only observable at sufficiently low temperatures.

1.5. *The year 1917: Einstein.

Einstein provides an intuitive derivation of Planck results by considering a system of harmonic oscillators with discrete energy levels E_n (possibly with some degeneracy Ω_n) in equilibrium with the radiation system originating from transitions between the states n,m and having the sought for energy density $u(\omega,T)$. Einstein starts with the number N_n of oscillators in the state n being given by the Boltzmann-Gibbs distribution P_n. These oscillators can decay into a state m with $E_m < E_n$ bei spontaneous emission of radiation of frequency $\hbar\omega = E_n - E_m$, the transition rate being

$$\frac{dN^s_{nm}}{dt} = N_n\cdot A_{nm}$$

A_{nm} is some coefficient that describes the process of spontaneous emission of radiation and was, at that time, unknown. On the other side, the existing radiation field can induce a transition from the state n to m with a rate

$$\frac{dN^i_{nm}}{dt} = u(\omega,T)\cdot N_n\cdot B_{nm}$$

Again, B_{nm} is some coefficient describing the process of interaction between harmonic oscillator and radiation field, at that time unknown. The state n, hovewer, can also be occupied when radiation of the frequency $\hbar\omega = E_n - E_m$ is absorbed by the system in state m with $E_n > E_m$, the absorption rate being

$$\frac{dN_{mn}}{dt} = u(\omega,T)N_m\cdot B_{mn}$$

Notice that, again, the coefficient B_{mn} is unknwon, but Einstein made the educated guess that $B_{nm} = B_{mn}$. We will prove the identity of the so called Einstein coefficients during the course of this lecture and we will compute them explicitly. In equilibrium we must have the relation

$$\frac{dN_{mn}}{dt} = \frac{dN^s_{nm}}{dt} + \frac{dN^i_{nm}}{dt}$$

which translates to

$$u(\omega, T)N_m \cdot B_{mn} = N_n \cdot A_{nm} + u(\omega, T) \cdot N_n \cdot B_{nm}$$

Using $P_{n,m} = \frac{N_{n,m}}{N}$, with N being the total number of oscillators we obtain

$$P_m \cdot u(\omega, T) \cdot B_{nm} = P_n \cdot A_{nm} + u(\omega, T) \cdot P_n \cdot B_{nm}$$

Solving for the sought for density $u(\omega, T)$ and taking into account that

$$\frac{P_n - P_m}{P_n} = 1 - e^{\frac{\hbar\omega}{k_B \cdot T}}$$

we obtain

$$u(\omega, T) = \frac{A_{nm}}{\hbar \cdot \omega \cdot B_{nm}} \cdot \frac{\hbar \cdot \omega}{e^{\frac{\hbar\omega}{k_B \cdot T}} - 1}$$

$u(\omega, T)$ consists of two factors: the one on the right hand side is similar to the average energy of a quantum mechanical harmonic oscillator with a frequency ω. This tells us that the radiation system also behaves as a gas of harmonic oscillators. The prefactor contains the ratio of two transition rates and must be further specified in order to find the result originally written down by Planck for $u(\omega, T)$. To determine the prefactor we use the strategy of considering the radiation ensemble, from the point of view of thermodynamics, as a set of independent harmonic oscillators and we count the number of such oscillator in a cavity of volume V. To count the number of radiation degrees of freedom we consider the radiation field in the volume V with a continuum of frequencies ω, so that we have to consider the eigenmodes of a cavity of size L^3 with the density of eigenmodes at a given frequency. The eigenmodes are characterized, under the assumption of proper boundary conditions for the radiation field within the cavity, by the $\vec{k}$ values

$$\vec{k} = \frac{\pi}{L}(n_x, n_y, n_z)$$

with $n_i = 0, \pm 1, \pm 2, ...$ The number of eigenmodes within a sphere of radius q in the $\vec{k}$ reciprocal space is given by $\frac{L^3}{3\pi^2} \cdot \frac{\omega^3}{c^3}$, using the relation $\omega = c \cdot q$. The density of states at ω amounts to the derivative of $\frac{L^3}{3\pi^2} \cdot \frac{\omega^3}{c^3}$ with respect to ω:

$$\frac{dN}{d\omega} = \frac{L^3}{\pi^2 c^3} \cdot \omega^2$$

Assigning the $k_B T$ equipartition value to any of these eigenmodes we obtain the Rayleigh-Jeans classical spectral density

$$u(\omega, T)\,|_{\text{classical}} = \frac{1}{\pi^2 c^3} \cdot \omega^2 \cdot k_B \cdot T$$

We require now that the high temperature limit of the expression

$$u(\omega, T) = \frac{A_{nm}}{\hbar \cdot \omega \cdot B_{nm}} \cdot \frac{\hbar \cdot \omega}{e^{\frac{\hbar\omega}{k_B \cdot T}} - 1}$$

be $u(\omega, T)\,|_{\text{classical}}$. This procedure fixes the pre-factor uniquely and produces the Planck result

$$u(\omega, T)\,|_{\text{Planck}} = \frac{1}{\pi^2 c^3} \cdot \omega^2 \cdot \frac{\hbar \cdot \omega}{e^{\frac{\hbar\omega}{k_B \cdot T}} - 1}$$

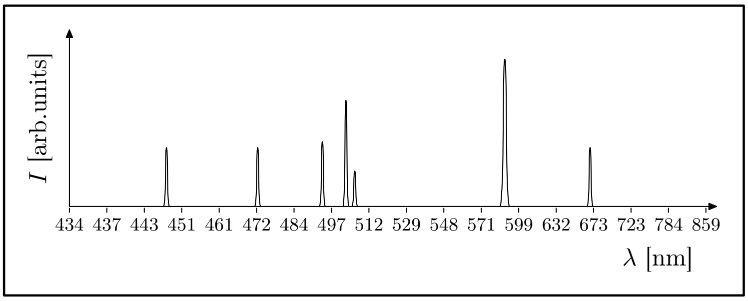
This expression also allows to determine a general law between the coefficients A_{nm} for spontaneous emission and B_{nm} for induced emission:

$$\frac{A_{nm}}{B_{nm}} = \frac{\hbar\omega_{nm}}{\pi^2 \cdot c^3}$$

which can be computed within the framework of advanced quantum mechanics.

1.6. The year 1913: Bohr

The most compelling experimental evidence for the existence of quantized energy levels were the emission spectrum of hydrogen and other atoms, which were measured accurately e.g. already at the end of the 19th century by optical spectroscopy. The light emitted by an ensemble of atoms was analyzed by a prism into its color-components, and only a set of well-defined colors were observed. Alternatively, one sends white light through a medium and observe the formation of well-defined black lines after analyzing the transmitted light according the its frequency.

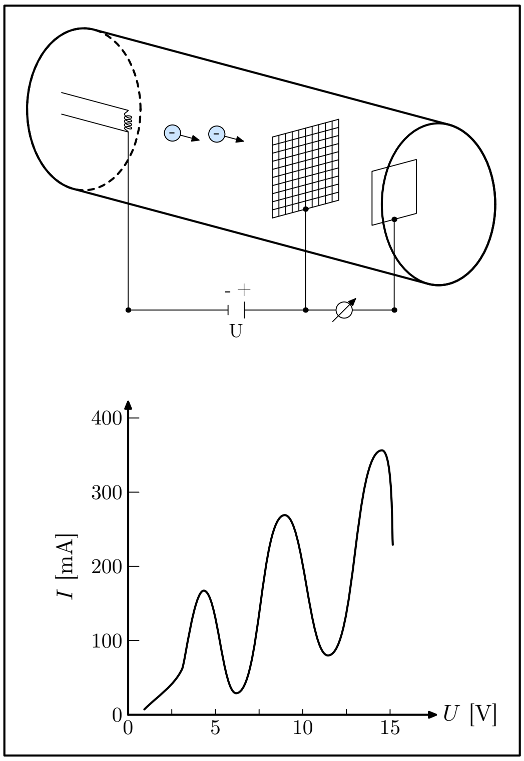


Within a tube, a gas – in this case He – is brought to glow. The light emitted is sent into a prisms and analyzed according to its colors. Characteristic peaks are observed in the intensity distribution, indicating that optical excitation in atoms occurs at well defined wave-lengths. This points to the presence of discrete energy levels in atoms that can absorb or emit light only with characteristic frequencies. See https://experimente.phys.ethz.ch/de/100/10002/20004/30559/

Of all atoms, atomic hydrogen emits the simples spectrum which was explained in 1885 by Balmer with the empirical formula

$$\omega_{nm} = R(\frac{1}{m^2} - \frac{1}{n^2})$$

where $n, m = 1, 2, \ldots\ n > m$ and the constant R (known as the Rydberg frequency) $R = 2\pi \cdot 3.310^{15} s^{-1}$, which was introduced to match the discrete emission lines observed by spectroscopy of the H-vapor. A further experimental outcome was the observation by Franck and Hertz between 1911 and 1914. In a vacuum tube, filled with typically mercury vapor, electrons are generated by means of a hot cathode and are accelerated by a controllable electrical voltage U between the emitting cathode and a grid. When U is increased the current grows, producing, at low voltages, a characteristic Ohm $I - V$-curve. The electrons collide with the Hg atoms and this provide an observable resistance. At some characteristic voltages (in Hg e.g. about 4.9 eV) the current drops sharply. This is interpreted as the kinetic energy of the electrons high enough to excite the Hg atom into a state Hg^*. By doing this excitation, the incoming electron looses its energy and ”vanishes” from the current. Evidently this excitation occurs at a well defined energy. The successive decay produces light with a well defined wavelength ($\lambda = 253.6 nm$ in Hg) corresponding to a far-ultraviolet light frequency. The outcome of the Franck and Hertz experiment is thus to be considered as a proof of discrete energy levels in an atom.



Top: Frank-Hertz experimental set-up. The glass tube is filled with e.g Hg vapor . An electron beam is accelerated by a variable voltage U. Most electrons are absorbed on the metal wires of the grid, but some pass through the grid and are collected at a collector. Their intensity is measured as a function of U. The collector has a negative voltage with respect to the grid, so that only electrons are detected which exceed a certain threshold for the kinetic energy. Bottom: typical current-voltage characteristic in a Franck-Hertz experiment. Because of inelastic collisions the electrons lose some of their kinetic energy, which is used to excite the atoms. Therefore, the inelastically scattered electrons are prevented from reaching the collector by the reverse voltage between the grid and the collector and are absent in the intensity measurement. This corresponds to a decrease of current in the $I-V$ characteristic. See https://experimente.phys.ethz.ch/de/100/10002/20004/30210/

The existence of quantized frequencies led Bohr to postulate the formation of quantized energy levels for the electronic structure of atoms. The simplest formula that would immediately explain the empirical Balmer law for the Hydrogen atom would be a set of discrete energy levels given by

$$E_n = -R_y \cdot \frac{1}{n^2}$$

with $R_y = \hbar \cdot R$ being what will be known as the unit of energy ”1 Rydberg” (about 13.6 eV).
The contribution of Bohr is the development of a simple, empirical, model that explains the quantization and compute R in terms of known natural and universal constants. He suggested to consider the electrons perfoming circular orbits around the nucleous (Rutherford model), at a well-defined energy E and with a well defined angular momentum L. The solution of the Newton problem of a central potential provides following equations for circular orbits:

$$E = -\frac{1}{4\pi\epsilon_0} \cdot \frac{e^2}{r} + \frac{L^2}{2mr^2} \qquad r = \frac{L^2 \cdot 4\pi\epsilon_0}{m \cdot e^2}$$

Eliminating r from these equations one obtains

$$E = -\frac{1}{2} \cdot \frac{e^4}{(4\pi\epsilon_0)^2} \cdot \frac{1}{L^2}$$

This result shows that a quantization of the energy levels in the Hydrogen atom is tantamount of imposing a quantum rule for the angular momentum of

$$L = n \cdot \hbar$$

with $n = 1, 2, 3, ...$ This produces the correct energy levels

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

Notice a very important element of the Bohr quantization rule: the ground state circular orbit has angular momentum $1 \cdot \hbar$. This result will be contradicted later by Schrödinger, which finds for the ground state of the Hydrogen atom an angular momentum of 0. From this result one can read out that the Rydberg energy:

$$\frac{1}{2} \cdot \frac{e^4 \cdot m}{(4\pi\epsilon_0)^2 \cdot \hbar^2}$$

is actually made up of physical constants such as the mass and the charge of the electron, $4\pi\epsilon_0$ and $\hbar$.

1.7. *The years 1913-1917: A. Sommerfeld.

A. Sommerfeld proposed that the quantization rules

$$E_n = n\hbar\Omega$$

for the harmonic oscillator and

$$L = n\hbar$$

for the angular momentum are special cases of a more general quantization rule. Consider one degree of freedom characterized by the spatial coordinate q and the momentum p. Consider e.g. the closed orbit $p(q)$ at a given energy resulting from the classical equation $\frac{p^2}{2m} + V(q) = E$ (one dimensional motion). Then Sommerfeld suggested the empirical rule of quantization for a classical closed orbit $p(q)$ as

$$\oint dqp(q) = n \cdot h$$

where the integral extends over a certain range of values q allowed for the orbit. Take, as example, the one-dimensional classic harmonic oscillator. Then energy conservation writes

$$\frac{p^2}{2m} + \frac{1}{2} m \cdot \Omega^2 \cdot q^2 = E$$

and the quantity $\int dq \cdot p$ is the area of the ellipse determined in the (p, q) coordinate-space by this equation. This means that

$$\oint pdq = \pi\sqrt{2mE} \cdot \sqrt{\frac{2E}{m}} \cdot \Omega^{-1} = \frac{2\pi \cdot E}{\Omega}$$

Sommerfeld quantization rule, applied to this results, reads

$$\frac{2\pi \cdot E}{\Omega} = n \cdot h$$

which is equivalent to Planck result

$$E = n \cdot \hbar \cdot \Omega$$

Sommerfeld quantization rule is important because it can be applied to a larger class of system that the ones investigated by Planck (one dimensional harmonic oscillator) and Bohr (circular orbits). For instance, using the quantization rule of Sommerfeld one can proceed to quantize elliptic orbits in the Kepler problem as well. When applied to systems with more that one degree of freedom, Sommerfeld quantization rule reads

$$\oint p_k dq_k = n_k \cdot h$$

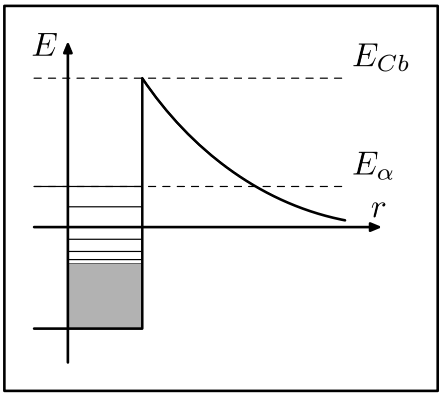
the index k labeling the corresponding degree of freedom. This generalization allows to introduce more quantum numbers – each referring to a separate degree of freedom. By means of the hypothesis of Sommerfeld, the electronic structure of atoms and the atomic spectra could be organized in a more systematic way and transitions between the various quantized states could be e.g. excluded because prohibited by some selection rules (see later in this lecture). Finally, the quantization rule of Sommerfeld set a very specific classical quantity at the center of more rigorous essays to develop a quantum mechanics from first priciples: the $\oint pdq$ which is known in classical mechanics as the ”action”. However, despite its generality and usefulness, the quantization rule of Sommerfeld is a tautology: it explains the quantization of some quantities by starting with a quantization. The Schrödinger equation is fundamentally different: is finds quantized quantities without imposing any quantization to start with.

1.8. 1921: Hendrika Johanna van Leeuwen.

Bohr, in his dissertation, and later Mrs. van Leeuwen in a rigorous paper, provided and exact proof of a very interesting result of statistical mechanics. If one assume that electrons in matter follow Newton mechanics, than the average magnetic moment arising from the electron system at finite temperatures is exactly vanishing. In other words, there is no magnetic matter. Yet the experimental facts about magnetism in solids are as old as the history of mankind: some materials have the property of producing a sizable magnetic field that attracts or repel other materials. One of the first references to the magnetic properties of what we know now to be magnetite Fe_3O_4 (lodestone) is by 6th century BCE Greek philosopher Thales of Miletus. The name ”magnet”may come from the lodestones found in Magnesia. In China, the earliest literary reference to magnetism lies in a 4th century BC book called Book of the Devil Valley Master: "The lodestone makes iron come or it attracts it". The lodestone based compass was used for navigation in medieval China by the 12th century. To explain the observation of magnetism in matter Ampere proposed, using Maxwell equations, that the origin of the magnetism in matter are the electrons that circulate at atomic levels and produce an atomic current, which, in turn, generates, via Maxwell equations, the magnetic field responsible for the magnetic behavior of matter. This model of magnetism is indeed correct: we now know with certainty that it is the finite angular momentum of electrons in atoms which generates the so-called atomic magnetic moment μ, which, in turn, produces the magnetic field. Bohr-van Leeuwen theorem is about the statistical average of the magnetic moment – a kind of equipartition theorem for magnetic moments –: at finite temperatures one must integrate over all states that can produce a magnetic moment in an atom, and this integration makes the average magnetic moment vanishing. We have seen that the equipartition theorem for the contribution of the oscillations to the specific heat in solids gives the temperature independent average Dulong-Petit value of k_B. The equipartition theorem, applied to the ensemble of magnetic moments, gives a vanishing value of the average magnetic moment at any temperature. We have learned from Einstein that the wrong result of classical physics can be bypassed by quantizing the energy of the mechanical motion. We will indeed learn that the Bohr-van Leeuwen catastrophe can be bypassed by quantizing the values of the angular momentum (Landau, 1930).

1.9. The years 1896-1900: Becquerel and M. Curie

The unstable atomic nuclei change spontaneously into other nuclei. They emit particles and very short-wave radiation. Natural radioactivity was discovered by Becquerel in 1896 and studied in depth by Pierre and Marie Curie. The emitted particles are either electrons or so-called α particles, i.e. He-cores. Later, Rutherford discovered the artificial nuclear transformation i.e. the ability, by shelling of particles, to transform nuclei of an element into nuclei of another element. Understanding the α decay of the heavy nuclei represents a challenge that is beyond the reach of classical physics.



Potential energy of a α particle as a function of the distance to the nucleus.

In fact, consider the reaction

$$U_{92}^{238} \longrightarrow Th_{90}^{234} + \alpha_2^4$$

where the emitted α particle has an energy of about 7.1 MeV (E_α in the figure). If one model the potential energy of an α particle within the nuclous by a potential well and outside it we allow the Coulomb law, one finds that the particle has to overcome a so-called Coulomb-barrier (E_{Cb} in the figure) in order to escape the nucleus. According to Coulomb's law, the Coulomb barrier amounts to $E_{Cb} = z_\alpha \cdot Z_K e^2/(4\pi\epsilon_0 R_K)$, with the core radius $R_K \approx 1.310^{-15} A^{1/3}$ (A: mass number)), Z_K the nuclear charge and z_α the charge number of the α particle. For a U_{92}^{238} decay one obtains, with $z_\alpha = 2$ and $Z_k = 92-2 = 90$ a barrier of $E_{Cb} = 32.1 MeV$. The ejected α particles have an energy of $7.1 MeV$. This means that the α particle originates from a level of the potential well which is below the Coulomb barrier. On its way out off the nucleus, the α particle in a radioactive decay must have traveled through a classically forbidden space region. The explanation of this one major failure of classical physics in term of tunnel effect by G. Gamov was one of the most convincing proofs of the validity of quantum mechanics at atomic level.

1.10. The year 1923: De Broglie

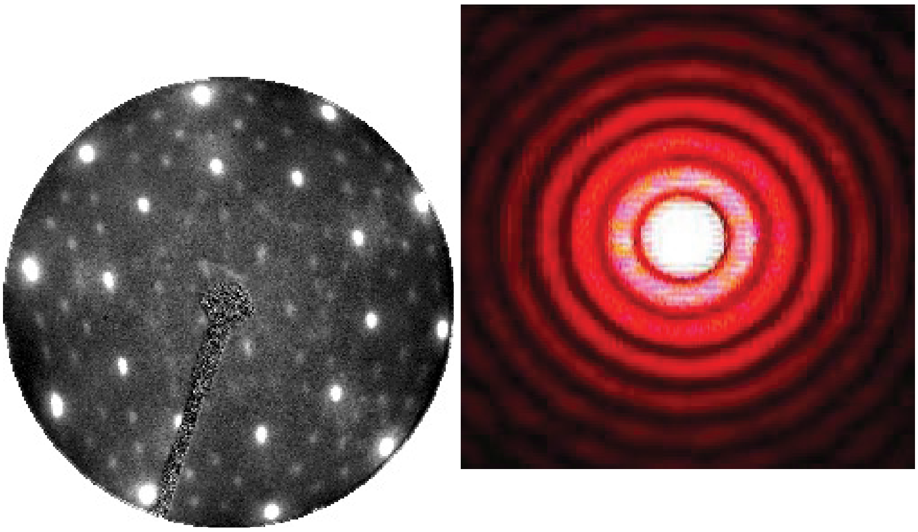
The photon hypothesis of Einstein foresees that electromagnetic waves consist of particles with energy $\hbar \cdot \omega$ and momentum $\hbar \cdot k$. With the photon hypothesis, Einstein established as system of particles that has the strange (for particles) property of undergoing e.g. diffraction from a small object. De Broglie, in his dissertation, translated the classical energy and classical momentum of a particle into its frequency and wavenumber:

$$\omega = \frac{E}{\hbar} \quad k = \frac{p}{\hbar}$$

thus assigning to ordinary matter a ”wave-like” character. De Broglie did not know what was the physical meaning of his hypothesis, although he proposed a possible consequence of it. He used the relation $\frac{\vec{p}^2}{2m} = E$ to compute the wavelength for a free particle of a given energy E and found (expressed with practical units)

$$\lambda[\text{in nm}] = \sqrt{\frac{1.54}{E[\text{in eV}]}}$$

Inserting $E = 100$ one obtains $\lambda \approx 0.1$ nm. Accordingly, De-Broglie proposed that one should expect to observe diffraction patterns when electrons are scattered e.g. at crystals (lattice constant: 0.1 nm). This prediction was indeed confirmed in 1927 (Davisson and Germer) by electron scattering at Ni crystal surfaces and was one of the most significant discoveries in the history of science. Diffraction was then observed with other particles, like neutrons. Recent experiments have observed diffraction even in atomic-and cluster- beam scattering experiments at artificial lattices (see e.g. J.P. Tonnies et al., Physikalische Blätter **56**, 53-55, 2000 and, for C_{70}, A. Zeilinger et al., Phys. Blätter, March 2000, p. 379)). The mathematical and physical aspects of diffraction will be the subject of an advanced topics chapter in this manuscript.



Left: Low-energy electron diffraction on a surface with tenfold symmetry (courtesy of M. Erbudak). The electrons are directed with about 100 eV kinetic energy onto the surface of a quasi-crystal and the backscattered electrons are visualized on a phosphorescent screen. Right: diffraction of red light at a small circular aperture. The laser beam arrives at a circular aperture with diameter comparable to its wavelength. The diffraction pattern recorded at a screen placed behind the aperture consists of concentric rings of intensity maxima separated by dark rings of intensity minima. `https://experimente.phys.ethz.ch/de/100/10002/20012/30164/`

2. Quantum mechanics 1.0: The Schrödinger equation.

2.1. The year 1926

(E. Schrödinger, Quantisierung als Eigenwertproblem I, Annalen der Physik 79 (1926), 361–376). The development of wave mechanics involved a lot of guessing and is therefore somewhat non-rigorous. We do not need to follow exactly the historical pathway, due originally to Schrödinger. His way of thinking was based on recognizing the action of classical mechanics as the central quantity and on assigning to it the unit $\hbar$ in order to generalize Hamilton-Jacobi mechanics to quantum mechanics. Here we prefer to follow an argument proposed a posteriori by M. Born.

• **The eigenvalue hypothesis.** The appearance of discrete energy levels reminds us of a similar discretization in classical physics. A classical model that contains discrete values as solutions is that of an oscillating string with boundary conditions. Let us recall the solution of this model in one spatial dimension, for mathematical simplicity. It starts with the amplitude of the oscillation (transversal to the string) being described by a scalar function $u(x,t)$. This scalar function obeys the wave equation

$$\frac{\partial^2 u(x,t)}{\partial t^2} = v^2 \cdot \frac{\partial^2 u(x,t)}{\partial x^2}$$

In classical physics this equation describes the propagation of a disturbance $u(x,t)$, like ripples on the surface of water, sound waves, electromagnetic waves and so on. v is the velocity of the disturbance in the medium and depends on the properties of the medium. Referring to light waves, we can think of v as $\frac{c}{\mu}$, c being the velocity of light in vacuum and μ the material-dependent refractive index. This equation is often formulated together with some boundary conditions. An example of boundary conditions is provided by the requirement that, at all times, $u(0,t) = u(L,t) = 0$, i.e. the amplitude is required to by vanishing at the ends of the interval $[0,L]$. We know from classical wave mechanics that these boundary conditions are referring to a situation of a vibrating string held fixed at two different points. We know that this physical system has well defined vibrational frequencies ω_n and we would like to find them by solving the wave equation. To find ω_n we insert the eigenmode Ansatz

$$u(x,t) = \psi(x) \cdot e^{-i \cdot \omega_n \cdot t}$$

into the wave equation, leading to the classical Helmholtz equation for the x-dependent part $\psi(x)$ and the sought-for frequencies ω_n:

$$\frac{\partial^2 \psi(x)}{\partial x^2} = -\frac{\omega_n^2}{v^2} \cdot \psi(x)$$

One can also write this equation using, on the right-hand side, the sought-for wave number k_n:

$$\frac{\partial^2 \psi(x)}{\partial x^2} = -k_n^2 \cdot \psi(x)$$

• **The eigenvalue problem.** We recall that, mathematically speaking, an equation were, on the left-hand side, one operates onto a sought-for function f with some kind of "device"(more precisely: an **"operator"**)[1] while on the right-hand side one has the sought-for function multiplied by a scalar λ, is referred to as an "eigenvalue" equation:

$$\mathscr{L} f = \lambda \cdot f$$

---

[1] In the case of the Helmholtz equation, the operator is the negative of the second derivative

DEFINITION. The eigenvalue equation AND the boundary conditions build the so-called **eigenvalue problem** for the sought-for eigenvalues λ_n and the sought-for eigenfunctions f_n, n being a (possibly discrete) number that is used to label the solutions.

DEFINITION. The set of values $\{\lambda_n\}$ that solve the eigenvalue problem builds the spectrum of the operator $\mathscr{L}$.

COMMENTS.

1. An eigenvalue equation has, in general, solutions for any value of λ, but if BOUNDARY CONDITIONS are introduced – such as keeping the spring fixed at some edges $\pm\frac{L}{2}$ ($\psi(\pm\frac{L}{2})=0$) – it has only solutions for a discrete set of values λ_n (the eigenfrequencies ω_n (or the eigen-wavenumbers k_n) in the case of the vibrating spring with fixed ends).

2. Schrödinger thought, in his seminal paper, that one should search for an equation similar to the Helmholtz equation, augmented with adequate boundary conditions, for finding quantized **energy** values.

• **The Schrödinger equation (SE).** The Helmholtz equation, as it is, does not contain $\hbar$. To introduce $\hbar$, we follow a method suggested later by M. Born. One starts from the Helmholtz equation

$$\frac{\partial^2\psi(x)}{\partial x^2}=-k_n^2\cdot\psi(x)$$

and then enforce De Broglie hypothesis for a particle with a potential energy $\Phi(x)$:

$$k_n^2=\frac{p^2}{\hbar^2}=\frac{2m}{\hbar^2}\cdot[E-\Phi(x)]$$

to write the "Helmholtz equation" (or time independent Schrödinger equation) for a particle (here generalized to three coordinates):

$$-\frac{\hbar^2}{2m}\triangle\psi(x,y,z)+\Phi(\mathbf{x})\cdot\psi(x,y,z)=E_n\cdot\psi(\mathbf{x})$$

COMMENTS.

1. Mathematically speaking, the SE consists, on the left-hand side, of an operator acting on the sought-for wave function. The right hand side contains the sought-for eigenvlaues multiplied by the sought-for eigenfunctions. Provided some suitable boundary conditions are attached to it, it should become an eigenvalue problem.

2. This equation is non-classical, as it contain $\hbar$ and one expects this equation to produce discrete "energy eigenvalues" E_n, provided some adequate boundary conditions are attached to the function $\psi(x,y,z)$ – just as the classical Helmoltz equation produces discrete "eigenfrequencies" when supplemented with suitable boundary conditions.

• **The boundary condition.** The finding of a suitable boundary condition is related to the interpretation of the sought-for wave functions. Schrödinger suggested to allow complex valued wave functions[2]. He than suggested to dispose of the electron as being point-like matter and to replace it with a continuous charge density $\rho(x)$ which he set to be the square of the absolute value of the complex wave **amplitude** $\psi(\mathbf{x})$:

$$\rho(\mathbf{x})=e\cdot\mid\psi(\mathbf{x})\mid^2$$

(e is the charge of the electron). The motive of using $\mid\psi(\mathbf{x})\mid^2$ as the physically relevant quantity for expressing the charge density can be found within a precise evaluation of the **classical Helmholtz** equation.

Given is a complex valued solution $\psi(\mathbf{x})$ of the **classical** Helmholtz equation

$$\triangle\psi(x,y,z)+\mathbf{k}^2(\mathbf{x})\cdot\psi(x,y,z)=0$$

---

[2] He knew that the solutions of the **classical** Helmholtz equation must be real valued functions: think that the amplitude of a wave propagating on the surface of a see is a measurable, real valued quantity.

CLAIM: The quantity

$$\mathbf{J} \doteq \frac{1}{2i}\left(\bar{\psi}(\mathbf{x})\nabla\psi(\mathbf{x}) - \psi(\mathbf{x})\nabla\bar{\psi}(\mathbf{x})\right)$$

has the property

$$\nabla \cdot \mathbf{J} = 0$$

PROOF: We write, without loss of generality,

$$\psi(x,y,z) \doteq A(\mathbf{x}) \cdot e^{i\cdot S(\mathbf{x})}$$

and show that

$$\nabla(A^2(\mathbf{x}) \cdot \nabla S(\mathbf{x}) = 0$$

For this purpose, we rewrite the Helmholtz equation in terms of $A(\mathbf{x})$ and $S(\mathbf{x})$. From

$$\nabla(A(\mathbf{x}) \cdot e^{i\cdot S(\mathbf{x})}) = \left(\vec{\nabla}A(\mathbf{x}) + i \cdot A(\mathbf{x})\nabla S(\mathbf{x})\right)e^{i\cdot S(\mathbf{x})} \quad \triangle = \nabla \cdot \nabla$$

we obtain

$$\begin{aligned} \triangle(A(\mathbf{x}) \cdot e^{i\cdot S(\mathbf{x})}) + \mathbf{k}^2(\mathbf{x}) \cdot (A(\mathbf{x}) \cdot e^{i\cdot S(\mathbf{x})}) &= \left([\triangle A(\mathbf{x}) - A(\mathbf{x}) \cdot (\nabla S(\mathbf{x}))^2]\right) \cdot e^{i\cdot S(\mathbf{x})} + \mathbf{k}^2(\mathbf{x}) \cdot (A(\mathbf{x}) \cdot e^{i\cdot S(\mathbf{x})}) \\ &+ i \cdot ([A(\mathbf{x})\triangle S(\mathbf{x}) + 2 \cdot \nabla A(\mathbf{x}) \cdot \nabla S(\mathbf{x})]) \cdot e^{i\cdot S(\mathbf{x})} \\ &= 0 \end{aligned}$$

The vanishing of the imaginary part produces the equation

$$A(\mathbf{x})\triangle S(\mathbf{x}) + 2 \cdot \nabla A(\mathbf{x}) \cdot \nabla S(\mathbf{x}) = 0$$

This equation writes also

$$\nabla \underbrace{\left(A^2(\mathbf{x}) \cdot \nabla S(\mathbf{x})\right)}_{\mathbf{J}} = 0$$

and this concludes the proof.

COMMENTS.

1. The vanishing of the divergence establishes the vector **J** as a stationary current density vector that propagates an excitation with density

$$A^2(\mathbf{x}) = \mid \psi(\mathbf{x}) \mid^2$$

2. It turned out later (see the discussion in the Chapter on "QM 2.0: the foundations") that Schrödinger did the right thing by using the square of the absolute value of the wave function as the physically relevant quantity, and not the wave function itself.

3. It, however, turned out (see the discussion in the Chapter on "QM 2.0: the foundations") that the assigning of the square of the absolute value of the wave funtion to a charge density is probably wrong, as recent experiments show (see the Nobel prize 2022). However, when dealing with the bound states of single-electron quantum mechanical problems in atoms, molecules and solids the Schrödinger interpretation is qunatitatively correct in predicting the actual charge distribution. We therefore will stick, for the time being, to Schrödinger interpretation.

4. Boundary conditions to the SE are, e.g., the continuity of the wave function and of its partial derivative, because only in this case the SE has a physical meaning. A further important boundary condition is originating from the charge density interpretation: the wave function must be square integrable, i.e.

$$\int_{\mathscr{R}^3} dV \mid \psi(x,y,z) \mid^2 < \infty$$

5. We summarize here some facts about the wave function for a single electron in Euclidean space.

- $(x,y,z) \doteq \mathbf{x}$ specify the possible coordinates that a particle can assume. Typically, they extend over the entire Euclidean space.
- $\psi(\mathbf{x})$ is called the "amplitude" of the "electron wave" and can be a complex function.

- $|\psi(\mathbf{x})|^2$ is called the "charge density" of the "electron wave".
- The integral

$$\int_{\mathscr{R}^3} |\psi(\mathbf{x})|^2 \, dV$$

must be finite as the electron has a finite"charge" and can be normalized to "1".

- In the literature one also finds

$$|\psi(\mathbf{x})|^2 \cdot dV$$

as the "charge" within a volume element dV.

6. As the SE is, mathematically speaking, related to the classical Helmholtz equation, it can generate the phenomena of diffraction predicted by De-Broglie and later indeed discovered using electrons in 1927 by Davisson and Germer and, independently, by G.P. Thomson. We will treat problems related to "scattering states" in a specific Chapter.

7. This equation was a great achievement of Schrödinger: he obtained the correct value for the discrete energy levels of hydrogen, containing the correct physical constants, starting from an equation which did not explicitly show neither the discreteness nor the value of the sought for energy levels – one speaks of such a computation as "starting from first-principles".[3]

8. In the literature one also finds the following, equivalent version of the SE:

$$\boxed{\triangle\psi(x,y,z) + \frac{2m}{\hbar^2}\cdot[E_n - \Phi(x,y,z)]\cdot\psi(x,y,z) = 0}$$

This version stresses the fact that the SE is a partial differential equation with coefficients that can depend on the variables (x,y,z) (non-constant coefficients).

The correspondence principle. Schrödinger proposed a simple translation key to guess quantum mechanics from classical mechanics This translation key goes under the name of "correspondence principle".

In classical mechanics, Newton equation can be, ultimately, obtained by a so called variational principle, applied to the **Hamilton function**. The Hamilton function expresses the energy of a particle as a function of its coordinates $\mathbf{x}, \mathbf{p}$:

$$H(\mathbf{x},\mathbf{p}) = \frac{\mathbf{p}^2}{2m} + \Phi(\mathbf{x})$$

He noticed that the operator acting (on the left hand side) onto the wave function in the equation

$$\left[-\frac{\hbar^2}{2m}\triangle + \Phi(\mathbf{x})\right]\psi(x,y,z) = E_n \cdot \psi(\mathbf{x})$$

can be simply obtained if one makes the following substitutions

$$\mathbf{x} \rightarrow \mathbf{x}$$
$$\mathbf{p} \rightarrow -i \cdot \hbar \cdot \nabla$$

in the classical Hamilton function.

COMMENTS.

1. Considering the Hamilton function of classical physics for a given system and using the correspondence principle will turn out to be a very useful starting point to formulate the SE for that system.

2. The more formal aspect of the correspondence principle is that it systematically introduces a set of

---

[3] (quote: "In dieser Mitteilung mochte ich zunächst an dem einfachsten Fall des (nichtrelativistischen und ungestörten) Wasserstoffatoms zeigen, dass die übliche Quantisierungsvorschrift sich durch eine andere Forderung ersetzen lässt, in der kein Wort von "ganzen Zahlen" mehr vorkommt. Vielmehr ergibt sich die Ganzzahligkeit auf dieselbe natürliche Art, wie etwa die Ganzzahligkeit der Knotenzahl einer schwingenden Saite".

mathematical entities – called ”operators” – that ”act” onto wave functions, such as

position operator	$\mathbf{x}$
momentum operator	$-i\cdot\hbar\cdot\nabla$
Hamilton operator	$\left[-\frac{\hbar^2}{2m}\triangle+\Phi(\mathbf{x})\right]$

In the Schrödinger quantum mechanics, operators are, typically, constructed using the coordinates or their derivatives.

3. In the operator language, the Schrödinger equation is **”the eigenvalue equation of the Hamilton operator”**.

4. In the Chapter ”Quantum mechanics 2.0: the foundations” we will learn more about the wider significance of operators in quantum mechanics and about possible ways to represent them.

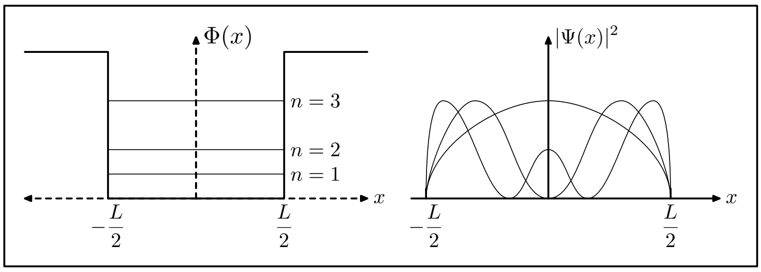
2.2. Applications

We leave the solution of the quantum mechanical hydrogen atom to a later chapter and instead provide here an application of the ideas by Schrödinger to mathematically simple but physically representative model systems.

2.2.1. One dimensional potential well

In this problem, we actually simulate an atom as providing a one-dimensional potential well for a particle with mass m.

$$\Phi(x)=\begin{cases}\Phi_0>0 & L/2\leq x\leq -L/2\\ 0 & -L/2\leq x\leq L/2\end{cases}$$



Left: one-dimensional potential well for a particle with mass m. The position of some bound states – labeled $n=1,2,3$ *is indicated. They will be computed in this section. Right: the ”charge” density for the bound state wave functions for* $n=1$*-state (knots at* $x=\pm\frac{L}{2}$*),* $n=2$*-state (knots at* $x=\pm\frac{L}{2},0$*) and* $n=3$*-state (knots at* $x=\pm\frac{L}{2},0,\pm\frac{1}{3}\frac{L}{2}$*). The bound state wave functions will also be computed explicitely in this section.*

Of course, this model is far away from the realistic case of an electron moving within the Coulomb field of a proton: the motion is a one-dimensional one, the depth of the well is finite and its shape is not similar to the Coulomb potential. Yet it provides a localized decrease of the potential energy which mimics the situation in a more realistic Coulomb potential. The configuration space (the range of the variables in the amplitude $\psi(x)$) is the entire x-coordinate. The Schrödinger equation (SE) for time independent – also called "stationary" – states of well-defined energy E reads

$$\frac{d^2}{dx^2}\psi(x)+\frac{2m}{\hbar^2}\cdot(E-\Phi(x))\psi(x)=0$$

It is a differential equation of the type

$$a_2\frac{d^2}{dx^2}f(x)+a_0(x)\cdot f(x)=0$$

$f(x)$ is the sought for solution, a_2, a_0 are coefficients. We know such differential equations from the physics of the classical harmonic oscillator, but there is a fundamental difference: the x-dependent potential introduces a dependence of at least one coefficient from the independent variable x (in this case a_0). The solution of such equations with non-constant coefficients is not trivial and we will learn more about this later. We therefore proceed with a simplified model. As the sought for quantities are the possible energies of the bound states of the particle, which we suspect to be close to the bottom of the potential well, we replace the problem with a similar but simpler one: we just set Φ_0 to infinity, hoping that this approximation will not affect too much those states which a closest to the bottom of the potential well. In the domain where $\Phi(x)=\infty$, we have to look for a solution for which the term $\Phi(x)\cdot\psi(x)$ in the Schrödinger equation remains finite, despite the potential diverging, so that the SE still makes sense. For $\Phi(x)\cdot\psi(x)$ to remain finite we choose $\psi(x)\equiv 0$ in the segments $L/2\le x\le -L/2$. In the interval $[-\frac{L}{2},\frac{L}{2}]$ the wave function may take on finite values. One important "boundary condition" is that the integral $\mid\psi(x)\mid^2$ be finite:

$$\int_{-L/2}^{L/2}\psi^*(x)\cdot\psi(x)\cdot dx=1$$

This is not the only boundary condition: $\psi(x)$ must be continuous – a physical "charge density" at one point cannot depend on whether we approach that point from right or from letf:

$$\psi(x)\mid_{\pm\frac{L}{2}}=0$$

for the solutions of SE in the interval $[-\frac{L}{2},\frac{L}{2}]$.

In the interval $[-\frac{L}{2},\frac{L}{2}]$ the SE reads

$$\left[\frac{d^2}{dx^2}+\underbrace{\frac{2mE}{\hbar^2}}_{\doteq k^2}\right]\psi(x)=0$$

The new sought for parameter $k\doteq\sqrt{\frac{2mE}{\hbar^2}}$ has been introduced for simplicity. The choice of the letter k is due to the fact that $\sqrt{\frac{2mE}{\hbar^2}}$ has the unity of a wave number.

The mathematical structure of the SE is the one typical of the classical harmonic oscillator (but only the mathematical structure, this is NOT a classical harmonic oscillator) and is familiar enough to be able to state that this differential equation has two linearly independent solutions:

$$\cos(k\cdot x)\quad\text{and}\quad\sin(k\cdot x)$$

At first glance, k can assume any real value and the set of $k's$ produces a non-countable set of linearly independent solutions. The general solutions are linear combinations of $\cos(k\cdot x)$ and $\sin(k\cdot x)$ and the integral over the absolute square in certainly finite: the square integrability of the solutions is therefore given. This specific boundary condition is, in the present instance, fulfilled for any k, i.e. it does not lead to any quantization.

The vanishing of the solutions at the edge of the interval is a further boundary condition.

a) Requiring that

$$\cos(k\cdot L/2)=0$$

restricts the allowed k to a discrete set given by

$$k_n = \frac{(2n+1)\cdot\pi}{L} \quad n = 0,1,2,...$$

(negative n-values do not produce further linearly independent solutions). The corresponding energy levels are QUANTIZED by this boundary condition:

$$E_n = \frac{\pi^2\hbar^2}{2mL^2}\cdot(2n+1)^2$$

The normalized wave functions read

$$\sqrt{\frac{2}{L}}\cos\frac{(2n+1)\pi}{L}\cdot x \quad n = 0,1,2,...$$

b) Requiring that

$$\sin(k\cdot L/2) = 0$$

restricts the allowed k to a discrete set given by

$$k_n = \frac{(2n)\cdot\pi}{L} \quad n = 0,1,2,...$$

(negative n-values do not produce further linearly independent solutions). The corresponding energy levels are QUANTIZED by this boundary condition:

$$E_n = \frac{\pi^2\hbar^2}{2mL^2}\cdot(2n)^2$$

The normalized wave functions read

$$\sqrt{\frac{2}{L}}\sin\frac{(2n)\pi}{L}\cdot x \quad n = 0,1,2,...$$

Notice that $n = 0$ is not allowed, as the corresponding wave function is identical to 0 and therefore cannot be normalized to 1.

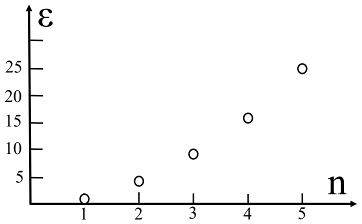
COMMENTS.

1. The solution of this problem produces a quantization of the energy values of the bound states of a particle in a potential well. The boundary conditions act to produce this quantization. The energy values are described by the expression

$$E_n = \frac{\pi^2\hbar^2}{2mL^2}\cdot(n)^2 \quad ; \quad n = 1,2,,3,....$$

2. The index n, used to label the energy levels, is called a ”quantum number”.

3. The energy values obtained by solving the SE with boundary conditions are usually represented by a graphic called ”term scheme” or ”electronic structure”.



Term scheme (electronic structure) for a particle in an infinite potential well. The energy (in units of $\frac{\pi^2\cdot\hbar^2}{2\cdot m\cdot L^2}$ is plotted as a function of the quantum number n.

4. When the wall of the potential well is ”soft” (i.e. not infinity) the wave function leaks into the forbidden segments $L/2 \leq x \leq -L/2$, but, as long as the states are close to the bottom of the well, this leakage does not modify essentially the solutions.

6. Sommerfeld quantization. One can obtain the quantized energy levels just computed by using Sommerfeld quantization rule. In classical mechanics, the motion of a mass moving in the interval $[-\frac{L}{2}, \frac{L}{2}]$ and bounded by impenetrable potential walls occurs at a given energy E, which can assume continuous values. The motion can be described with a graph in a p versus x diagram. p describes the momentum. The graph corresponds, in the present case, to a rectangle with horizontal side L and vertical height $\pm p$, where $p = \sqrt{2mE}$. The classical path of a particle is along the sides of this rectangle. The px- parameter space is called in classical mechanics the **phase space** (the "phase space" is not to be confused with the "configuration space", describing the set of possible values of the variable x). Classical mechanics foresees the possibility for the particle to assume continuous values for x (the spatial coordinate). This element is maintained in quantum mechanics. p assumes in classical mechanics also continuous values. From this situation one can move to quantum mechanics using Sommerfeld quantization rule. For this particular system the quantization rule requires that the area of a rectangle, when its lateral sides are run across clockwise, must be an integer multiple of h:

$$2p \cdot L = n \cdot h \quad \Leftrightarrow \quad p = \frac{n \cdot \pi}{L} \cdot \hbar$$

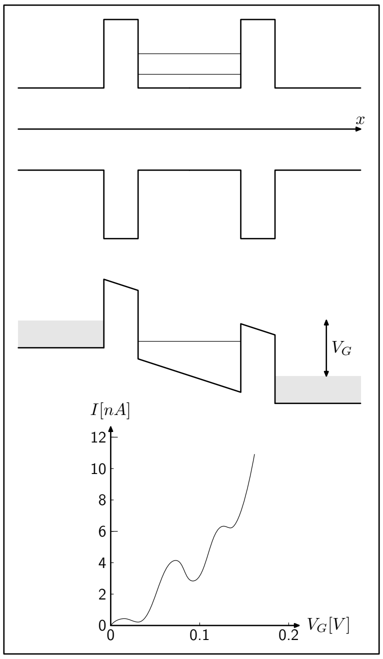
Sommerfeld quantization rule therefore selects a discrete number of rectangles in phase space. We can translate the quantized momentum to a quantized energy by using the classical translation key $p = \sqrt{2mE}$ and we obtain, for the possible energy values of the quantum mechanical motion, exactly the result that we have computed by solving the SE with boundary conditions:

$$E_n = \frac{\pi^2 \hbar^2}{2mL^2} \cdot n^2 \quad ; \quad n = 0, 1, 2,, 3, ...$$

A small but significant detail: Sommerfeld quantization rule allows the energy level with $n = 0$, which we have found to be excluded from the solutions of the SE.

Practical application: QUANTUM WELL

Modern technology has found important applications of this solution in the form of "artificial atoms". The most spectacular application of a particle in the box comes probably from semiconductor technology. A semiconductor, such as GaAs, has an energy gap in which stationary states of the electrons cannot exist. The energy gap is about 1 eV: an electron must overcome this barrier in order to contribute to current transport. When some of the Ga atoms are substituted by Al the energy gap gets bigger. If two layers of GaAlAs include a layer of pure GaAs, a potential box in the conduction band in the direction of stacking appears: a so-called quantum well is built.



Top: energy of the conduction band (top) and valence band (bottom) as a function of the stacking coordinate x. AlGaAs barrier layers a few nanometer thick bracket a GaAs layer of comparable thickness to form a quantum potential well for electrons. The potential well energy states are indicated by horizontal lines. Middle: The potential well is biased with application of a gate voltage V_G and the energy levels are bent. Bottom: Plot of a drain to source current as a function of the voltage V_G. The three peaks indicate resonant tunnelling (resonant tunnelling diode, Esaki, Chang and Tsu, IBM).

The height of the potential well is about 0.3 eV. The steepness of the potential barrier depends on how abrupt is the Al doping. Molecular Beam Epitaxy allows the concentration of the Al to drop to zero within an atomic location. That's why the steepness of the potential well can be very large and one obtain a box with walls that resemble a jump (even if not with infinite potential jump). One can also, if necessary, let the Al concentration decrease slowly, making a parabolic potential well in which harmonic vibrations are possible (an artificial quantum harmonic oscillator). For the quantization to be effective, the distance between adjacent energy levels in the potential well must be greater than, for example, their width. Scattering on crystal defects and/or on phonons are responsible for the broadening of discrete energy levels: the width at room temperature is at least 0.03 eV (corresponding to $k_B \cdot 300K$). From the inequality

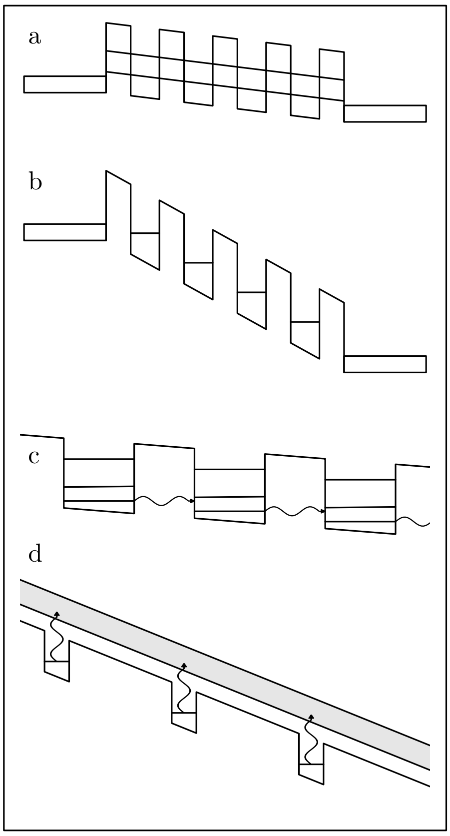
$$0.03eV < \pi^2\hbar^2/(2mL^2)$$

one obtains a maximum length of $L \leq 4$ nm for the quantization of energy levels in the well to be observable. This length represents a great challenge to the accuracy of MBE growth but is now achieved as a standard in routine semiconductor based devices.

Once the quantized levels have been established, one can bias the junction with a certain voltage across the quantum well between source (cathode, left) and anode (right, drain). At some bias voltage the quantized levels in the box match the energy levels of the source. If the potential barrier is sufficiently thin, source electrons can pass through the potential barrier via discrete energy levels in the box by quantum mechanical tunnelling (the phenomenon of tunnelling through a classically forbidden energy range, classically prohibited because in that region the particle has a negative kinetic energy, is a pure quantum mechanical effect and will treated later as a possible solution of the SE) and thereby reaching the anode. The presence of discrete energy levels offers an additional channel for electrons to flow through the well structure. The generation of additional current is only possible if the levels in the structure are properly positioned: one speaks of a **resonant tunnel diode**. The current-voltage characteristic goes through a series of maxima, which occur only at the resonant voltages, and a clear sign of energy quantization in the box. This non-linear current-voltage characteristic can be used, for example, as an amplifier or switch. There is a possibility that the electrons in the box loose energy by emitting light: thus one can also build a light-emitting diode, which emit light only at certain voltages.

SUPERLATTICE.

A super-lattice with many periodic potential wells opens new possibilities. The potential barriers must still be sufficiently thin, so that the electrons can easily tunnel. In a) the voltage is small, the energy levels in the box are practically at the same height: According to quantum mechanics, a one-dimensional artificial band structure is formed (analogous to the phonon bands): a one dimensional crystal appears. Typically, electrons can flow across the super lattice, forming a kind of artificial metal. In b), no band is formed because the slope is too strong and the levels do not overlap: an insulator builds up. Such artificial super-lattices are an ideal tool to drive a metal-insulator transition by applying a voltage. In c) the voltage is chosen so that in each case the 1st level coincides with the 2nd of the following box.



Electron transport in a super lattice. a): Conduction-band profile showing alternating layers of two semiconductor materials of different band-gap width. The darker area represents an allowed energy mini-band arising from the super-lattice periodicity. Its slope is due to an applied voltage bias. b): A stronger bias field causes enough slope to misalign the energy levels in successive wells. As a result, the mini band is replaced by a Wannier-Stark ladder of discrete states. c): In a super-lattice with weak coupling between the wells one can increase the field until the lowest level in each well is aligned with one of the excited states in the next well. This produces electron transport with an alternating sequence of tunnelling and energy relaxation. This scheme,with electrons tunnelling into the second excited state of a neighboring well, forms the basis for a all solid state infrared laser. d): Quantum well infrared detector. Electrons from the wells are excited by the incident photons into the extended states above the barriers and transported to the contacts.

The electron can only traverse the structure by emitting light. The emission of light takes place between discrete levels, so that the emitted light is monochromatic. If one insert the structure between reflecting mirrors, an infrared laser is produced, which is pumped by the application of voltage. In modern communication technology such solid state lasers are often used. The layer structure can also work as an infrared detector (d in the Figure).

2.2.2. Periodic boundary conditions

Consider a particle moving along the x-axis with vanishing potential. Assume that the wave functions are subject to periodic boundary conditions at the edge of the interval $[-\frac{L}{2},\frac{L}{2}]$:

$$\psi(-\frac{L}{2}) = \psi(+\frac{L}{2})$$

The SE reads

$$\left[\frac{d^2}{dx^2} + \underbrace{\frac{2mE}{\hbar^2}}_{\doteq k^2}\right]\psi(x) = 0$$

We find a solution of this equation using the Ansatz $\psi(x) = e^{\lambda\cdot x}$ with λ to be determined so that the SE is fulfilled. Inserting this Ansatz we obtain the algebraic equation

$$\lambda^2 + k^2 = 0$$

One possible solution is $\lambda = i\cdot k$, so that a possible wave function solving the SE is $e^{i\cdot k\cdot x}$. The boundary condition, applied to this solution, reads:

$$e^{i\cdot k\cdot \frac{L}{2}} = e^{-i\cdot k\cdot \frac{L}{2}} \Longleftrightarrow e^{i\cdot k\cdot L} = 1$$

The values of k solving this last equation can be expressed as

$$k_n = \frac{2\pi}{L}\cdot n; \quad n \in \mathscr{Z}$$

The possible values of the energy are quantized and amount to

$$E_n = \frac{\hbar^2}{2m}\cdot(\frac{2\pi}{L})^2\cdot n^2$$

The solutions of the SE can be normalized to "1": one has to find the normalization constant a so that

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} dx\cdot \mid a\mid^2 \cdot \mid e^{i\cdot k_n\cdot x}\mid^2 = 1$$

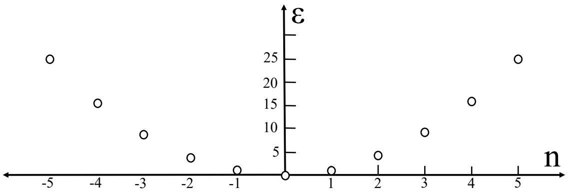
One possible value for a is $\sqrt{\frac{1}{L}}$. This leads to an infinite set of numerable wave functions

$$\{u_n(x) = \frac{1}{\sqrt{L}}\cdot e^{i\frac{2\pi\cdot n}{L}\cdot x}\} \quad ; \quad n = 0, \pm 1, \pm 2, \ldots$$

The index n is a quantum number used to numerate the discrete energy levels and the set of corresponding wave functions.

COMMENT.

We observe here (see the next figure) that to a given energy level there are two linearly independent eigenfunctions. This multiplicity is known in the terminology of quantum mechanics as "degeneracy" of the eigenvalue E_n. One speks of **"degenerate" eigenvalue**. We will encounter degenerate eigenvalues in more complex problems.



Term scheme (electronic structure) for a free particle with periodic boundary conditions. The energy (in units of $\frac{\pi^2\cdot 4\cdot\hbar^2}{2\cdot m\cdot L^2}$ is plotted as a function of the quantum number n.

2.2.3. The QM Harmonic Oscillator

We consider a particle of mass m moving in a one-dimensional space under the influence of the potential energy

$$\Phi(x) = \frac{K}{2}x^2 + \phi_0 \doteq \frac{m\omega^2}{2}x^2 + \phi_0$$

K is a "material" parameter used to express the strength of the "spring" at which the mass is attached. In an atomic system, K is related to the second derivative of the potential energy between two next-neighbour constituents of the material. In the literature of both the classical and quantum mechanical harmonic oscillator K is often transformed into a characteristic frequency $\omega^2 \doteq K/m$ for convenience. The SE reads

$$\frac{d^2\varphi(x)}{dx^2} + \frac{2m}{\hbar^2}(E - \frac{m\omega^2}{2}x^2 - \phi_0)\varphi(x) = 0$$

E are the sought for energies. The boundary condition is the square integrability of the sought for wave functions:

$$\int_{-\infty}^{+\infty} \mid \varphi(x) \mid^2 dx < \infty$$

We will show that the fulfillment of this boundary condition leads to a quantization of the energy states of a harmonic oscillator, as originally postulated by Planck. We will find that

$$E_n = \hbar\omega(n + \frac{1}{2}) + \phi_0$$

with $n = 0, 1, 2, \ldots$. The wave functions to the energy values E_n are given by

$$\varphi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) e^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2}$$

The $H_n(\xi)$ are the Hermite polynomials:

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

also

$$H_0(\xi) = 1$$
$$H_1(\xi) = 2\xi$$
$$H_2(\xi) = 4\xi^2 - 2$$
$$H_3(\xi) = 8\xi^3 - 12\xi$$
$$H_4(\xi) = 16\xi^4 - 48\xi^2 + 12$$

and so on. The energy of the ground state is not 0 but has a higher value: it is a " zero point energy ", which represents a "dithering" (quantum fluctuation). We recall e.g. the wave function of the ground state:

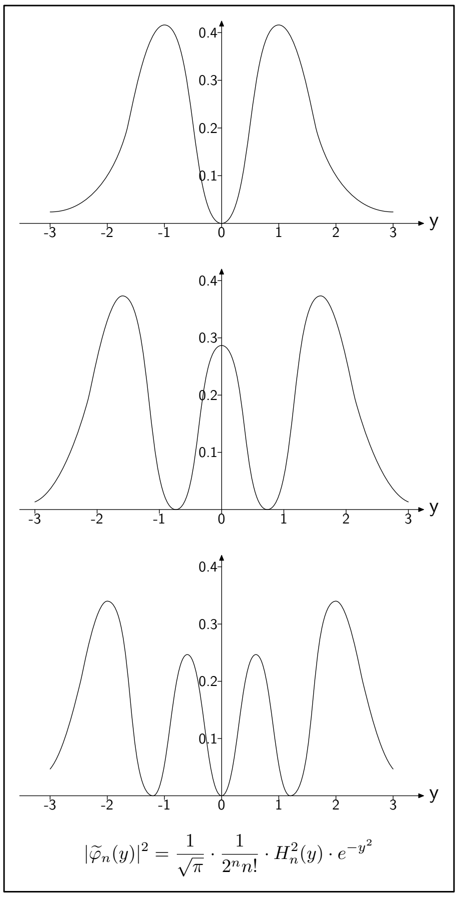
$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} e^{-\frac{1}{2}\frac{m\omega}{\hbar}x^2}$$

This is a Gaussian centered at $x = 0$. In the wave functions one recognizes a characteristic length

$$\xi \doteq \sqrt{\frac{\hbar}{m\omega}}$$

For a Silicon atom vibrating in a lattice with a characteristic phonon frequency in the terahertz regime this characteristic length, over which the square of the absolute value of the wave function has a sizeable value, is about 0.01 nm. The characteristic length can be used to plot the (dimensionless) square of the wave functions as a function of the dimensionless quantity $y \doteq \frac{x}{\xi}$ (from top to bottom: the probability density for the $n = 1, 2, 3$ states as a function of the position).

$$\mid \tilde{\varphi}(y) \mid^2 = \frac{1}{\sqrt{\pi}} \cdot \frac{1}{2^n \cdot n!} \cdot H_n^2(y) e^{-y^2}$$



$$|\widetilde{\varphi}_n(y)|^2 = \frac{1}{\sqrt{\pi}} \cdot \frac{1}{2^n n!} \cdot H_n^2(y) \cdot e^{-y^2}$$

From top to bottom: the probability density for the $n = 1,2,3$ *states as a function of the position.*

*Details of the mathematical solution

We will discuss the details of the mathematical solution, with the aim of showing the link between boundary conditions and quantization. We start from the equation

$$\frac{d^2\varphi(x)}{dx^2} + \frac{2m}{\hbar^2}(E - \frac{m\omega^2}{2}x^2 - \phi_0)\varphi(x) = 0$$

and carry out the variable substitution

$$\frac{2(E-\phi_0)}{\hbar\omega} \equiv \epsilon \quad \xi = x\sqrt{\frac{m\omega}{\hbar}}$$

We obtain the ordinary differential equation of the second degree

$$\frac{d^2\varphi}{d\xi^2} + (\epsilon - \xi^2)\varphi = 0$$

For the wave function to be normalizable, $\varphi(\xi)$ must decay with ξ for $\xi \to \pm\infty$. We use therefore the Ansatz $\varphi(\xi) = v(\xi) \cdot e^{-\frac{\xi^2}{2}}$ and find the equation for the function $v(\xi)$

$$v'' - 2\xi v' + (\epsilon - 1)v = 0$$

(the dash means derivative with respect to ξ). Its solution is based on the general theory of ordinary differential equations with non-constant coefficients: one try to find a solution using a series of powers of the variable.

$$v(\xi)=\sum_{\nu=0}^{\infty} c_\nu \xi^\nu$$

Inserting into the DE one obtains

$$\sum_{\nu=0}^{\infty}[c_{\nu+2}(\nu+2)(\nu+1)+c_\nu(-2\nu+\epsilon-1))]\xi^\nu=0$$

This is a sum of powers of ξ which can only vanish if the coefficient of each power vanishes:

$$\frac{c_{\nu+2}}{c_\nu}=\frac{2\nu+1-\epsilon}{(\nu+1)(\nu+2)}$$

With the help of this recursion formula one can start by setting c_0 or c_1 and determine all coefficients of the power series. By starting with c_0 a solution $\varphi(\xi)$ is produced, which contains only even powers of ξ. These solutions are even with respect to change of the sign of ξ. Those solutions that start from c_1 are therefore uneven.

It is particularly important now to check whether the exponential factor multiplied with the power series provide solutions which are really normalizable, because only when solutions of the DE can be normalized they can also be regarded as physically meaningful. For the normalizability, the behaviour of the solutions at infinity is decisive, so that we have to study the asymptotic behavior of the power series. For large ν the recursion formula goes to

$$\frac{c_{\nu+2}}{c_\nu}\rightarrow 2/\nu$$

This asymptotic behaviour has the same recursion formula of the coefficients of

$$e^{2\xi^2}=\sum_{\mu=0}^{\infty}\xi^{2\mu}\frac{2^\mu}{\mu!}$$

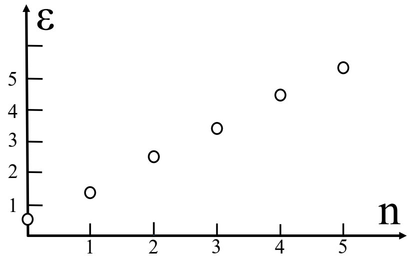
The infinite series thus behaves asymptotically as $e^{2\xi^2}$. When considered with the factor $e^{-\xi^2/2}$, this asymptotic behaviour destroy the sought for quadratic integrability of $\varphi(\xi)$. To avoid this "catastrophe" one must find a way of terminating the series at a certain finite $\nu=n$. A polynomial results, which does no longer repesents a "danger" for the square integrability of the solutions. Let us then terminate the series by requiring that $c_{n+2}=0$, where n can be $0,1,2,....$. The equation $c_{n+2}=0$ is equivalent to requiring

$$2n+1-\epsilon=0$$

But this last equation is physically very relevant: it imposes the condition that ϵ (i.e. the energy) can only assume discrete values, namely $\epsilon=2n+1$, $n=0,1,....$ This condition produces a **quantization** of the allowed energy levels of a linear harmonic oscillator:

$$E_n=\hbar\omega(n+\frac{1}{2})+\phi_0$$

numerated by a quantum number $n=0,1,2,....$



Term scheme of the one dimensional quantum oscillator. $\epsilon = \frac{E-\phi_0}{\hbar\omega}$ *is plotted versus the quantum number n.*

3. The Hydrogen atom

3.1. Introduction

Any fundamentally new theory, to be accepted at all, had to deal, in particular, with the discrete energy levels of the H-atom and explain the resulting spectroscopic laws. This is because the experimental evidence of atomic spectra about their discreteness and their exact values were absolutely certain – and inexplicable on the basis of classical physics. E. Schrödinger saw this task as his main concern in his first communication (1926) on the quantization of the eigenvalue problem. The solution proposed by Bohr

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

($n = 1, 2, 3, \ldots$) on the basis of an empirical quantization rule of the angular momentum was, in fact, correct, but it was based on an empirical fact, i.e. suggested ad hoc to explain the specific experimental result represented by Balmer formula. The scope of a fundamentally new theory was to obtain the same results but from first principles – such as starting from the wave equation and, possibly, some boundary conditions. Schrödinger did exactly this and in doing so created a fundamentally new theory of nature.

3.2. General symmetry aspects of the solution

The Schrödinger equation for a mass in the field of force represented by a **spherically symmetric** potential $V(\mathbf{r}) = V(x^2 + y^2 + z^2)$ reads

$$[-\frac{\hbar^2}{2m} \triangle + \Phi(x^2 + y^2 + z^2)]\psi(\mathbf{r}) = E\psi(\mathbf{r})$$

E are the sought for eigenvalues. The model of the atom underlying this equation is one where the proton – which is much heavier than the electron – is assumed to have an infinite mass and reside at the origin of the coordinate system. In this way, one eliminates from the problem the degrees of freedom represented by the coordinates of the proton. The boundary condition reads

$$\int dV \cdot |\psi(\mathbf{r})|^2 < \infty$$

We will show that requiring the square integrability of the eigenfunctions will finaly lead to the quantization of the energy eigenvalues and to the formula empirically proposed by Bohr.
The spherical symmetry of $\Phi(\mathbf{x})$ is best considered with the choice of the spherical coordinates (r, ϑ, φ):

$$x = r\sin\vartheta\cos\varphi \quad y = r\sin\vartheta\sin\varphi \quad z = r\cos\vartheta$$

$$r \in [0, \infty[\quad \vartheta \in [0, \pi] \quad \varphi \in [0, 2\pi]$$

The inverse transformations read

$$r = \sqrt{x^2 + y^2 + z^2} \quad \vartheta = \arccos(\frac{z}{r}) \quad \varphi = \arctan(\frac{y}{x})$$

In order to transform the Laplace operator, which includes derivatives, we need to learn how to write expressions like $\frac{\partial}{\partial x}$ in spherical coordinates. We use the chain rule to write

$$\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} = \begin{pmatrix} \partial r/\partial x = \sin\vartheta\cos\varphi & \partial\vartheta/\partial x = \frac{1}{r}\cdot\cos\vartheta\cos\varphi & \partial\varphi/\partial x = -\frac{\sin\varphi}{r\cdot\sin\vartheta} \\ \partial r/\partial y = \sin\vartheta\sin\varphi & \partial\vartheta/\partial y = \frac{1}{r}\cdot\cos\vartheta\sin\varphi & \partial\varphi/\partial y = \frac{\cos\varphi}{r\cdot\sin\vartheta} \\ \partial r/\partial z = \cos\vartheta & \partial\vartheta/\partial z = -\frac{1}{r}\sin\vartheta & \partial\varphi/\partial z = 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial\vartheta} \\ \frac{\partial}{\partial\varphi} \end{pmatrix}$$

The transformation matrix for the partial derivatives is called the Jacobi matrix. We write then

$$\frac{\partial^2}{\partial x^2} = \sin\vartheta\cos\varphi\frac{\partial}{\partial r}[\frac{\partial}{\partial x}] + \frac{1}{r}\cos\vartheta\cos\varphi\frac{\partial}{\partial\vartheta}[\frac{\partial}{\partial x}] - \frac{\sin\varphi}{r\cdot\sin\vartheta}\frac{\partial}{\partial\varphi}[\frac{\partial}{\partial x}]$$

$$\frac{\partial^2}{\partial y^2} = \sin\vartheta\sin\varphi\frac{\partial}{\partial r}[\frac{\partial}{\partial y}] + \frac{1}{r}\cdot\cos\vartheta\sin\varphi\frac{\partial}{\partial\vartheta}[\frac{\partial}{\partial y}] + \frac{\cos\varphi}{r\cdot\sin\vartheta}\frac{\partial}{\partial\varphi}[\frac{\partial}{\partial y}]$$

$$\frac{\partial^2}{\partial z^2} = \cos\vartheta\frac{\partial}{\partial r}[\frac{\partial}{\partial z}] - \frac{1}{r}\sin\vartheta\frac{\partial}{\partial\vartheta}[\frac{\partial}{\partial z}] + 0\cdot\frac{\partial}{\partial\varphi}[\frac{\partial}{\partial z}]$$

Inserting in this second set of equations the transformation rules for $[\frac{\partial}{\partial x}]$, $\frac{\partial}{\partial y}]$ and $[\frac{\partial}{\partial z}]$, we obtain finally (after a lengthy computation)

$$\triangle = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial}{\partial r}) + \frac{1}{r^2}\Lambda$$

with

$$\Lambda \equiv [\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}(\sin\vartheta\frac{\partial}{\partial\vartheta}) + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}]$$

The SE in spherical coordinates reads:

$$\left[-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial}{\partial r}) - \frac{\hbar^2}{2mr^2}\Lambda + \Phi(r)\right]\psi(\mathbf{r}) = E\psi(\mathbf{r})$$

This is an homogeneous partial differential equation with some symmetries: the potential depends only on one of the variables (the radius r). One could also think, technically speaking, of a partial differential equation with boundary conditions specified along a surface (or a line, in two dimensions) where one coordinate is constant (e.g. the surface of a sphere in spherical coordinates or a circle in polar coordinates). In such "symmetric situations" the method of choice for solving the partial differential equation is the so-called separation Ansatz, where a solution is sought for consisting of the product of function of one independent variable only:

$$\psi(r,\vartheta,\varphi) = f(r)\cdot Y(\vartheta,\varphi)$$

The rationale of this Ansatz will become clear soon. The SE writes now (let us introduce the notation $\frac{\partial}{\partial r} \equiv'$ for simplicity):

$$(r^2 f(r)')' Y(\vartheta,\varphi) + f(r)\cdot\Lambda Y(\vartheta,\varphi) + \frac{2m}{\hbar^2}\cdot r^2[E-\Phi(r)]f(r)\cdot Y(\vartheta,\varphi) = 0$$

Let us divide by $f(r)\cdot Y(\vartheta,\varphi)$:

$$\frac{(r^2 f(r)')'}{f(r)} + \frac{2m}{\hbar^2}\cdot r^2[E-\Phi(r)] = -\frac{\Lambda Y(\vartheta,\varphi)}{Y(\vartheta,\varphi)}$$

By means of the separation Ansatz we have achieved an interesting result: the SE writes as an equation between a left-hand side, which depends only on the variable r (for this it is important that the potential only depends on r) and a right-hand side which only depends on the angle variables φ,ϑ. As this equation must hold for any value on the variables, we have

$$\frac{(r^2 f'(r))'}{f(r)} + \frac{2m}{\hbar^2}\cdot r^2[E-\Phi(r)] = \lambda \quad -\frac{\Lambda Y(\vartheta,\varphi)}{Y(\vartheta,\varphi)} = \lambda$$

where λ is some constant to be determined and is called the separation constant. By means of the separation Ansatz, we have achieved a separation of the problem into two equations: one for the variable r and one for the variables φ,ϑ. These two separate problems are often easier to solve than the original one. The radial equation reads

$$(r^2 f'(r))' + \frac{2m}{\hbar^2} \cdot r^2[E - \Phi(r)]f(r) - \lambda \cdot f(r) = 0$$

It contains the sought for eigenvalues E and the yet to be determined separation constant λ and must be solved with the boundary conditions of the square integrability of $f(r)$:

$$\int_0^\infty dr \cdot r^2 \cdot |f(r)|^2 < \infty$$

Notice that the SE contains parts which have different singular behaviour at the origin. We must require that the wave function be finite at the origin. With this requirement, we can achieve that singularities at the origin cancel out (the left-hand side of the SE must be equal zero everywhere, including the origin). The square integrability is therefore supplemented with the boundary condition that

$$f(0) = \text{const.}$$

The second problem contains the angular variables ϑ, φ, i.e. it is defined on the sphere:

$$\Lambda Y(\vartheta, \varphi) = -\lambda Y(\vartheta, \varphi)$$

We recognize here that the separation constant $-\lambda$ is the solution of an eigenvalue problem for the operator Λ, as the eigenvalue equation on the sphere comes about with the boundary conditions

$$Y(\varphi + 2\pi, \vartheta) = Y(\varphi, \vartheta) \quad Y(\varphi, \vartheta + 2\pi) = Y(\varphi, \vartheta)$$

COMMENT.

Solving the Laplace equation

$$\triangle u(x, y, z) = 0$$

or the Helmholtz equation

$$\triangle u(x, y, z) + k^2 u(x, y, z) = 0$$

with a separation Ansatz leads to the same eigenvalue problem on the sphere. This is the reason why the eigenvalue problem on the sphere is discussed e.g. in problems of electrostatics.

3.2.1. The eigenvalue problem on the unit sphere

The solution of the eigenvalue problem on the sphere is worked out in textbooks of mathematical physics. Here we quote those results that are relevant for finding the sought for eigenvalues E. It turns out that the boundary conditions narrow down the values of $-\lambda$ to a discrete set:

$$-\lambda_l = -l \cdot (l+1); \quad l = 0, 1, 2, 3, ...$$

One also find that each of these eigenvalues is exactly $2l+1$-times degenerate: The eigenspace E_l contains the $2l+1$ linearly independent spherical harmonics $Y_l^m(\vartheta, \varphi)$ $(m = l, l-1, ..., -l)$:

$$\{Y_l^m\} = \begin{cases} Y_l^l(\vartheta, \varphi) \\ Y_l^{l-1}(\vartheta, \varphi) \\ \cdot \\ \cdot \\ \cdot \\ Y_l^{-l}(\vartheta, \varphi) \end{cases}$$

The spherical harmonics are given for positive $m = 0, 1, ...l$ by

$$Y_l^m(\vartheta, \varphi) = \frac{(-1)^{l+m}}{2^l \cdot l!} \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} \cdot P_l^m(\cos\vartheta) e^{im\varphi}$$

and

$$Y_l^{-m} = (-1)^m \overline{Y_l^m}$$

The associated Legendre functions $P_l^m(\cos\vartheta)$ are defined in terms of the derivative of the Legendre polynomials:

$$P_l^m(\cos\vartheta) = (-1)^m(1-\cos^2\vartheta)^{\frac{m}{2}}\frac{d^m}{d\cos\vartheta^m}P_l(\cos\vartheta)$$

$$P_l(\cos\vartheta) = \frac{1}{l!\cdot 2^l}\frac{d^l}{d(\cos\vartheta)^l}[(\cos\vartheta^2-1)]^l$$

The spherical harmonics fulfill the normalization and orthogonality relations

$$\int \overline{Y_{l,m}}(\vartheta,\varphi)Y_{l',m'}(\vartheta,\varphi)\sin\vartheta d\vartheta d\varphi = \delta_{ll'}\delta_{mm'}$$

We give now a list of some useful spherical harmonics (from E.U. Condon, G.H. Shortley, "Theory of atomic spectra", Cambridge University Press, 1953).

$$Y_0^0 = \frac{1}{\sqrt{4\pi}}$$

$$\begin{cases} Y_1^1 = -\sqrt{\frac{3}{8\pi}}\sin\vartheta e^{i\varphi} \\ Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\vartheta \\ Y_1^{-1} = \sqrt{\frac{3}{8\pi}}\sin\vartheta e^{-i\varphi} \end{cases}$$

$$\begin{cases} Y_2^2 = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\vartheta e^{2i\varphi} \\ Y_2^1 = -\sqrt{\frac{15}{8\pi}}\sin\vartheta\cos\vartheta e^{i\varphi} \\ Y_2^0 = \sqrt{\frac{5}{4\pi}}(\frac{3}{2}\cos^2\vartheta - \frac{1}{2}) \\ Y_2^{-1} = \sqrt{\frac{15}{8\pi}}\sin\vartheta\cos\vartheta e^{-i\varphi} \\ Y_2^2 = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\vartheta e^{-2i\varphi} \end{cases}$$

By solving the eigenvalue problem on the sphere we have generated a set of possible values for the separation constant. These values will be now plugged into the radial eigenvalue problem, which will be solved to find the sought for energy eigenvalues. For a given separation constant – specified by the number l – there will be, possibly, an energy eigenvalue that we can designate with E_l, as its value depends on the integer l. In addition, we already know that the eigenspace to E_l will be at least $2l+1$ degenerate, and its eigenfunctions will be of the type

$$\{f_{l,E}(r)\cdot Y_l^m(\vartheta,\varphi)\} = \begin{cases} f_{l,E}(r)\cdot Y_l^l(\vartheta,\varphi) \\ f_{l,E}(r)\cdot Y_l^{l-1}(\vartheta,\varphi) \\ \cdot \\ \cdot \\ \cdot \\ f_{l,E}(r)\cdot Y_l^{-l}(\vartheta,\varphi) \end{cases}$$

The indexes (l,m) used to label the basis functions in the eigenspace E_l provide quantum numbers that characterize the energy level (E,l).

At this point, before we tackle the radial problem, it is important to discuss the physical meaning of the quantum numbers l,m and the eigenvalues $-\lambda_l$. We recall that classical orbits in a spherically symmetric potential were characterized by some conserved total energy, indicated by the parameter E, and the conserved component L_z of the angular momentum. These parameters assume, in classical physics, continuous values. In quantum mechanics the energy is used to classify the eigenstates – with the difference that, as shown by Bohr empirically, it assumes discrete values. The conserved angular momentum was also shown by Bohr to play a key role for determining the energy values. We therefore set up to search for the fingerprint of the angular momentum in the radial and angular problems. We find immediately this fingerprint after using the correspondence principle to translate the classical angular momentum to a quantum mechanical angular momentum:

$$\mathbf{L} \doteq \mathbf{r} \times \mathbf{p}$$

to

$$\mathbf{L} = -i\hbar[\mathbf{r} \times \nabla]$$

i.e.

$$-i\hbar \cdot \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ x & y & z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{vmatrix} = -i\hbar \begin{pmatrix} y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \\ z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \\ x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \end{pmatrix}$$

We now compute this operator in spherical coordinates. We compute, e.g.,

$$\begin{aligned} L_z &= -i\hbar \cdot [x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}] \\ &= -i\hbar \left\{ r\sin\vartheta\cos\varphi \left[\frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial \vartheta}{\partial y}\frac{\partial}{\partial \vartheta} + \frac{\partial \varphi}{\partial y}\frac{\partial}{\partial \varphi} \right] \right. \\ &= \left. -r\sin\vartheta\sin\varphi \left[\frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \vartheta}{\partial x}\frac{\partial}{\partial \vartheta} + \frac{\partial \varphi}{\partial x}\frac{\partial}{\partial \varphi} \right] \right\} \end{aligned}$$

Using again to the Jacobi matrix we obtain

$$L_z = -i\hbar\frac{\partial}{\partial \varphi}$$

The x and y components write

$$L_x = i\hbar \cdot \left(\sin\varphi\frac{\partial}{\partial \vartheta} + \cot\vartheta\cos\varphi\frac{\partial}{\partial \varphi} \right)$$

$$L_y = -i\hbar \cdot \left(\cos\varphi\frac{\partial}{\partial \vartheta} - \cot\vartheta\sin\varphi\frac{\partial}{\partial \varphi} \right)$$

We can finally compute $\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2$ and find

$$\mathbf{L}^2 = -\hbar^2 \left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial \vartheta}(\sin\vartheta\frac{\partial}{\partial \vartheta} + \frac{1}{\sin\vartheta}\frac{\partial^2}{\partial \vartheta^2} \right]$$

We recall that

$$\Lambda = [\frac{1}{\sin\vartheta}\frac{\partial}{\partial \vartheta}(\sin\vartheta\frac{\partial}{\partial \vartheta}) + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial \varphi^2}]$$

from which we can read out the relation between the Λ operator and $\mathbf{L}^2$

$$\mathbf{L}^2 = -\hbar^2 \cdot \Lambda$$

Because of this relation, we can make following statements:

- The $2l+1$-degenerate spaces E_l are eigenspaces of $\mathbf{L}^2$
- The within E_l, $\mathbf{L}^2$ has the eigenvalues

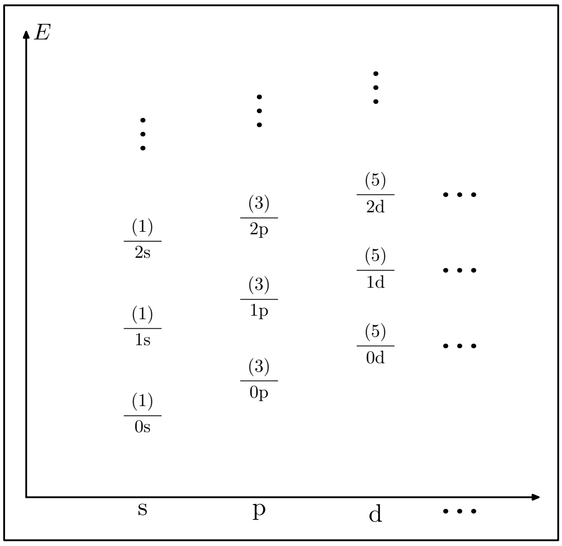
  $$\hbar^2 \cdot l \cdot (l+1)$$

  Henceforth we call l the ”orbital momentum quantum number”.
- The equation

  $$\underbrace{-i\hbar\frac{\partial}{\partial \varphi}}_{L_z} Y_l^m(\vartheta, \varphi) = \hbar \cdot m \cdot Y_l^m(\vartheta, \varphi); \quad m = -l, -l+1, ..., l$$

  shows that $Y_l^m(\vartheta, \varphi)$ is an eigenfunction of L_z with eigenvalue $\hbar \cdot m$. Henceforth we call m the ”magnetic quantum number” (the name ”magnetic” originates from the fact that m plays a special role in magnetic phenomena).

- When a particle is in a state within E_l one speak of the particle having orbital angular momentum l. l itself is NOT the eigenvalue of any operator.
- as
$$\sqrt{l(l+1)} > l$$
one can think of the orbital angular momentum never being exactly along the z-axis.
- For the states for the different values $l = 0, 1, 2, ...$ one often use the terminology that designates them with small Latin letters s, p, d, f, ... and so on. For example, the states with the orbital angular momentum zero ($l = 0$) are s-states, the states with $l = 1$ are called p-states, and so on.
- We are now ready to draw a provisional term-scheme for the possible energy states of a system in a spherical symmetric potential. We suspect that, for a given l – i.e. for a given separation constant – a set of energy levels will be produced by solving the radial equation, these energy levels being labeled by a further quantum number that we call n_r as it refers to the solutions of the radial SE.



A general term scheme for a quantum system with spherical symmetric potential. The vertical scale is the energy, the horizontal scale the orbital quantum number. The bulk of the figure contains possible energy levels which are supposed to form a discrete set. For a given value of l they are labeled by a quantum number n_r, which together with l is used to lebel the energy level: $E_{n_r,l}$. In one dimensional problems the orbital momentum quantum number is absent. The degeneracy of each level is indicated. The degeneracy is absent in one dimensional problems. For the specific case of a spherical symmetric potential the degeneracy is said to be essential, i.e. is dictated by the spherical symmetry and cannot be removed by modifying the potential – except if the potential is modified to depart from spherical symmetry. The relative position of the levels must be computed by solving explicitely the radial eigenvalue problem. We also do not know, at this point, whether the discrete set of energy levels continues to infinity or whether it stops at some value. We also do not know, at this point, whether there are eigenvalues which occupy a continuous set. All these matters are the results of the radial SE.

- The energy levels displayed in the figure have an important property: they are degenerate with respect to the quantum number m. Degeneracy is much rare in one dimensional problems, as the symmetry elements of a potential in one-dimension are limited to the identity and the inversion. For the specific case of a spherical symmetric potential the degeneracy with respect to m is said to be essential, i.e. is dictated by the spherical symmetry and cannot be removed by modifying the potential – except if the potential is modified to depart from spherical symmetry.

3.2.2. Solution of the radial problem

We now proceed toward solving the radial SE. As we know that the space spanned by $\{f_{E,l}(r) \cdot Y_l^m(\vartheta,\varphi)\}$ is an eigenspace of the Hamilton operator, we limit ourselves to solve the SE within this eigenspace. The SE within this eigenspace reads:

$$\left[\frac{-\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial}{\partial r})+\Phi(r)+\frac{\hbar^2\cdot l\cdot(l+1)}{2mr^2}\right]f_{E,l}(r)=E\cdot f_{E,l}(r)$$

Technical aspects

We now proceed to write the SE in a mathematically simple shape. We first set $f_{E,l} \doteq \frac{R_{E,l}}{r}$ and require that

$$R_{E,l}(r=0)=0$$

The SE for $R_{E,l}(r)$ reads

$$-\frac{\hbar^2}{2m}\frac{d^2R_{E,l}}{dr^2}+[\phi(r)+\frac{\hbar^2 l(l+1)}{2mr^2}]R_{E,l}=E_l\cdot R_{E,l}$$

This equation cannot be solved, in general, exactly: only for some special forms of the potential $\phi(r)$ and exact solution exists. One of these potentials is the Coulomb potential of an electron in the field of Z protons, i.e.

$$\phi(r)=-Ze^2/(4\pi\epsilon_0 r)$$

$Z=1$ describes the H-atom, $Z=2,3,..$ describes "hydrogen like" ions such as He^+, Li^{++}, etc., which contain only one electron. The SE now reads

$$\frac{d^2R_{E,l}}{dr^2}+[\frac{2mE}{\hbar^2}+\frac{2mZe^2}{4\pi\epsilon_0\hbar^2 r}-\frac{l(l+1)}{r^2}]R_{E,l}=0$$

It is convenient to introduce now the characteristic atomic quantities such as

$$a=\frac{4\pi\epsilon_0\hbar^2}{me^2}\approx 0.529\cdot 10^{-10}m$$

(the so called Bohr radius) and

$$E_a=\frac{e^2}{4\pi\epsilon_0 a}=\frac{me^4}{(4\pi\epsilon_0)^2\hbar^2}=27.21eV\doteq 1Ha$$

(a unit of energy called "Hartree"). We rewrite then the SE in terms of the dimensionless quantities

$$\rho\doteq r/a \quad \varepsilon=E/E_a$$

to obtain

$$[\frac{d^2}{d\rho^2}+2\cdot\varepsilon+\frac{2Z}{\rho}-\frac{l(l+1)}{\rho^2}]R(\rho)=0$$

Case a

If we think of the classical situation, we can conjecture that the souhgt-for bound states belong to a range of energies $2\varepsilon<0$. We therefore seek the solution of the SE in a situation where

$$2\varepsilon<0$$

To make this instance explicit, we set $2\varepsilon\doteq-\alpha^2$ and obtain

$$[\frac{d^2}{d\rho^2}-\alpha^2+\frac{2Z}{\rho}-\frac{l(l+1)}{\rho^2}]R(\rho)=0$$

- To gain a physically reasonable Ansatz we investigate the behaviour of this equation at large values of ρ. For $\rho \to \infty$ one can neglect the two last summands: the asymptotic equation reads

$$[\frac{d^2}{d\rho^2} - \alpha^2] R(\rho) = 0$$

The asymptotic solution ($\rho \to \infty$) reads

$$R(\rho) \to A \cdot e^{-\alpha\rho} + B \cdot e^{\alpha\rho}$$

As the wave function must vanish in the limit $\rho \to \infty$ we set $B = 0$.

- In order to determine the asymptotic behaviour of $R_{E,l}(\rho)$ for small ρ we insert the Ansatz $R_{E,l}(\rho) \sim \rho^{\gamma}$ into the SE, $\gamma \geq 0$ to be determined. In the lowest order in ρ the SE reads

$$[\gamma \cdot (\gamma + 1) - l(l+1)]\rho^{-\gamma-2} + ... = 0$$

As this equation must hold for any ρ we must set the coefficient to vanish. This gives a conditional equation for γ:

$$[\gamma \cdot (\gamma + 1) - l(l+1)] = 0$$

The solutions are

$$\gamma = -l$$

or

$$\gamma = (l+1)$$

The solution $\gamma = -l$ violates the condition that $R_{E,l}(\rho)$ must vanish at $\rho = 0$. This leaves us with a unique solution for γ.

- We put now an Ansatz forward that takes the asymptotic behaviour into account:

$$R_{E,l}(\rho) = e^{-\alpha\rho} \cdot \rho^{l+1} \cdot \sum_{k=0}^{\infty} a_k \rho^k$$

When we insert this Ansatz into the SE, we obtain that an infinite sum of powers of ρ must vanish for any ρ. This requires the vanishing of the coefficient of any power of ρ individually. The coefficient to the k-power contains a term proportional to a_{k+1} and a term proportional to a_k. Setting the coefficient to zero produces a recursion formula for the coefficients:

$$a_{k+1} = a_k \frac{2[\alpha(k+l+l) - Z]}{(k+l+2)(k+l+1) - l(l+1)}$$

For large k, the recursion formula crosses over to

$$\frac{\beta_{k+1}}{\beta_k} \to \frac{2\alpha}{k}$$

Such a recursion formula applies for the function

$$e^{2\alpha\rho}$$

If we do not limit the series expansion to a finite number of terms, then we get a non-square integrable function from the exponential growth for $\rho \to \infty$. We therefore require that the power series stops at some integer $k = n_r$ which can have the values $0, 1, 2, 3.....$ This requirement can be implemented with the equation

$$\alpha(n_r + l + 1) - Z = 0$$

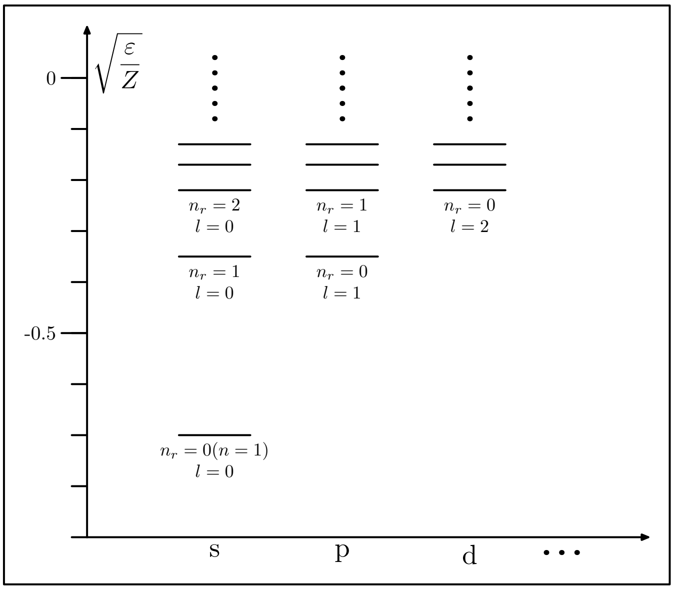
i.e.

$$\alpha = \frac{Z}{n_r + l + 1}$$

Recalling the relation between α and the dimensionless energy ε we obtain that the requirement of square integrability limits the energy to a discrete set given by

$$E_{n_r,l} = -\frac{me^4 Z^2}{2(4\pi\epsilon_0)^2\hbar^2} \cdot \frac{1}{(n_r + l + 1)^2}; \quad n_r = 0, 1, 2, 3, ...; \quad l = 0, 1, 2, 3, ...$$

This result contains the new quantum numbers n_r which labels the discrete energy eigenvalues, in addition to l. The term-scheme of the hydrogen, based on this exact result, is given in the figure.



Term-scheme of hydrogen like atoms as resulting from the solution of the radial equation. In the literature, the levels are often labelled by the quantum number $n = n_r + l + 1$ rather than by (n_r, l). This is confusing as it masks the fact that in more general spherical symmetric potentials ***two*** *separate quantum numbers are necessary to label the eigenvalues.*

In the literature, one uses the quantum number $n \doteq n_r + l + 1$ to label the energy levels, so that the contact with the empirical formula by Bohr

$$E_n = -\frac{me^4 Z^2}{2(4\pi\epsilon_0)^2 \hbar^2 \cdot n^2}$$

is made explicitely. n is known as the main quantum number. For $n_r = l = 0$ $(n = 1)$, e.g, we have the ground $1s$-state at -13.6 eV, i.e. one Rydberg $(1Ry = 1/2Ha)$. Its distance from $E = 0$ is the ionization energy of an electron in H-atom. Further states are found at $l = 0$ an $n_r = 1, 2, 3, 4, ...$ (see the figure caption for further discussion and n-assignment). One notices in the term-scheme that some states with different quantum numbers n_r and l are actually degenerate (for a given n, those states with $l = 0, 1, 2, ..., n-1$ are degenerate) i.e. they have the same energy. We point out that this degeneracy is so called an ”accidental” degeneracy: one can show that a small change away from the exact Coulomb potential (e.g. produced by screening, see later in atomic physics) removes the accidental degeneracy. The essential $2l+1$-degeneracy is, of course, maintained.

The limiting of the series to finite terms introduces characteristic polynomials in the build-up of the eigenfunctions: the Laguerre polynomials. The eigenfunctions of Z atoms, normalized to 1 read:

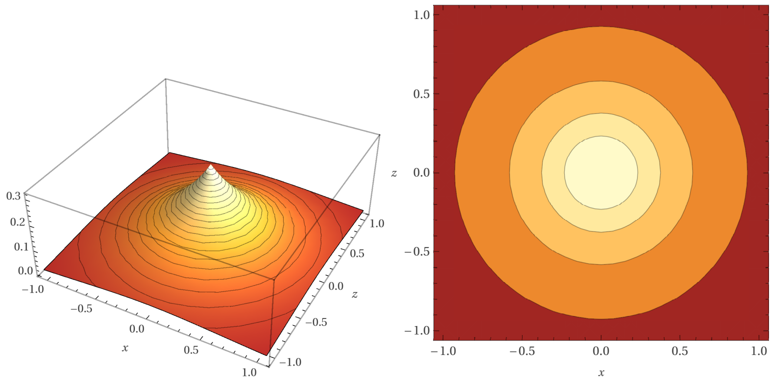
$$\psi_{n,l,m}(r,\vartheta,\varphi) = (a)^{-\frac{3}{2}} \cdot N_{nl} \cdot (\frac{2Zr}{a \cdot n})^l \cdot F\left(l+1-n, 2l+2, \frac{2Zr}{a \cdot n}\right) \cdot e^{-\frac{2Zr}{a \cdot n}} \cdot Y_n^m(\vartheta, \varphi)$$

$$N_{nl} = \frac{1}{(2l+1)!} \cdot \sqrt{\frac{(n+l)!}{2n \cdot (n-l-1)!}} \cdot (\frac{2Zr}{a \cdot n})^{\frac{3}{2}}$$

is a normalization factor. $F(a, c, z)$ is the confluent hyper-geometric function. It is typically defined as a infinite power series

$$F(\alpha,\beta,z)=1+\frac{\alpha}{\beta}\frac{z}{1!}+\frac{\alpha(\alpha+1)}{\beta(\beta+1)}\cdot\frac{z^2}{2!}+\ldots$$

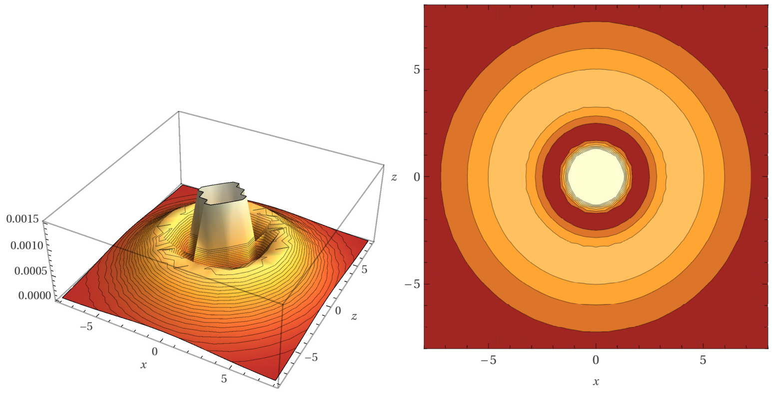
For $\alpha=0,-1,-2,\ldots$ $F(\alpha,\beta,z)$ is a polynomial. Some hydrogen-like orbitals are given here and sketched in the figures.



Computed by Wolfram|Alpha

Computed by Wolfram|Alpha

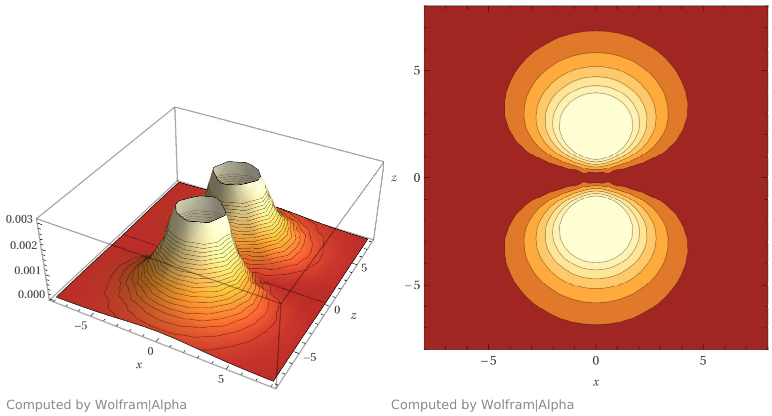
Left: Plot of the square of the $1s$*-wave function* $\psi(r,\vartheta,\varphi)_{1s}=\frac{2}{a^{3/2}}\cdot e^{-\frac{r}{a}}\cdot Y_0^0(\vartheta,\varphi)$ *in the* $x-z$*-plane.* a *is set to 1. Right: Contour plot.*



Computed by Wolfram|Alpha

Computed by Wolfram|Alpha

Left: Plot of the square of the $2s$*-wave function* $\psi(r,\vartheta,\varphi)_{2s}=\frac{2}{(2a)^{3/2}}\cdot e^{-\frac{r}{2a}}\cdot(1-\frac{r}{2a})\cdot Y_0^0(\vartheta,\varphi)$ *in the* $x-z$*-plane.* a *is set to 1. Right: Contour plot.*



Computed by Wolfram|Alpha

Computed by Wolfram|Alpha

Left: Plot of the square of the $2p_z$*-wave function* $\psi(r,\vartheta,\varphi)_{2p} = \frac{2}{\sqrt{3}(2a)^{3/2}} \cdot e^{-\frac{r}{2a}} \cdot (\frac{r}{2a}) \cdot Y_1^0(\vartheta,\varphi)$ *in the* $x-z$*-plane. a is set to 1. Right: Contour plot.*

Case b

We go back to the radial SE in the form

$$[\frac{d^2}{d\rho^2} + 2\cdot\varepsilon + \frac{2Z}{\rho} - \frac{l(l+1)}{\rho^2}]R(\rho) = 0$$

In classical Newton physics, open orbits appear in the range $\varepsilon \geq 0$. We consider therefore the range $\varepsilon \geq 0$ in the radial SE. Considering only the leading terms for $\rho \to \infty$ the radial SE in the range $\varepsilon \geq 0$ reads:

$$[\frac{d^2}{d\rho^2} + 2\cdot\varepsilon]R(\rho) = 0$$

The general solution is

$$R(\rho, \rho \to \infty) \approx A\cdot e^{i\cdot\sqrt{2\varepsilon}} + B\cdot e^{-i\cdot\sqrt{2\varepsilon}}$$

$R(\rho, \rho \to \infty)$ remains finite for any value of $\varepsilon \geq 0$ and the requirement of square integrability cannot be enforced. Accordingly, we do not have any tool to discretize the energy values: in the range $\varepsilon \geq 0$ we expect a continuum spectrum. The corresponding eigenstates are called ”scattering states”. They are orthogonal to the bound states appearing in the discrete part of the spectrum and build and essential component of the space of states of the SE for a radial potential. They are not square integrable. We will learn how to deal with such non-square integrable scattering states in a separate Chapter.

Part II.

Foundations of quantum mechanics

4. Formal aspects of the Schrödinger QM and QM 2.0.

4.1. Introduction

Schrödinger quantum mechanics introduces wave functions, operators acting on them and the eigenvalues of these operators, resulting from the solution of an eigenvalue problem. J. von Neumann, in his fundamental work ”Mathematische Grundlagen der Quantenmechanik”, Springer, Berlin, 1932 established a rigorous general mathematical framework for hosting the new elements entailed by Schrödinger quantum mechanics. The formulation of this framework is one scope of this Chapter. We will find that this mathematical framework is the key to the transition from Schrödinger quantum mechanics to an entirely equivalent formulation of quantum mechanics, based on using vectors to represent ”wave functions”, matrices to represent ”operators” and the eigenvalue problem of a matrix to represent the eigenvalue problem of an operator. The matrix formulation is due, originally, to Heisenberg. It has the advantage that problems of quantum mechanics can be cast and solved in situations where the wave functions and the operators cannot be represented with standard spatial variables, for instance situations involving the ”spin” state of the electron or the ”color” state of a quark. A second scope of this Chapter, accordingly, is to get aquainted with the matrix formulation of quantum mechanics. These are the ”formal” aspects dealth with in this chapter. Finally, the third topic of this Chapter: the finding of the ”observables” and their link to the outcome of an experiment. This third topic (in particular the ”jump” postulate by Dirac) we call QM 2.0, as it represents a genuine development with respect to QM 1.0 (the Schrödinger Quantum mechanics). QM 2.0 is currently much debated, because experiments aimed at verifying such fundamental questions have been recently designed and performed. Our approach will be one of establishing a fundamental understanding of the topic rather than discussing the recent advances.

4.1.1. The year 1926: M. Born

The proposal of interpreting the square of the absolute value of the wave function as a charge or particle density and thus of substituting point like particles with some continuous distribution of matter was controversial at the time Schödinger proposed it (and at later times as well). We quote, at this point, verbatim from the Nobel lecture by M. Born (1954):
”Schrödinger thought that his wave theory made it possible to return to deterministic classical physics. He proposed (and he has recently emphasized his proposal anew’s), to dispense with the particle representation entirely, and instead of speaking of electrons as particles, to consider them as a continuous density distribution (or electric density $e\cdot|\psi|^2$). To us in Göttingen this interpretation seemed unacceptable in face of well-established experimental facts. At that time it was already possible to count particles by means of scintillations or with a Geiger counter, and to photograph their tracks with the aid of a Wilson cloud chamber.”......”Again an idea of Einstein’s gave me the lead. He had tried to make the duality of particles - light quanta or photons - and waves comprehensible by interpreting the square of the optical wave amplitudes as probability density for the occurrence of photons. This concept could at once be carried over to the ψ-function: $|\psi(\mathbf{x})|^2$ ought to represent the **probability density** for electrons (or other particles).

This is the most celebrated ”statistical interpretation of the wave function” by M. Born.
As pointed out by Born, bound states are not suited to distinguishing the statistical character of the wave function, but scattering states such as those found in diffraction, are more telling. The diffraction pattern recorded e.g. in electron or atom scattering consists of registered ”bips” along a well-defined channel at a certain angle. The ”bips” mark the arrival of the particles. The number of ”bips” registered depends on the scattering angle, and the distribution of the number of registered ”bips” as a function of scattering angle follow the characteristic diffraction pattern consisting of maxima and minima. In other words: atoms and electrons are particles that have a certain probability to appear at some geometrical point. The statement by Born is that this probability is related to $|\psi(x,y,z)|^2$.

The statistical interpretation by Born produces a well-defined mathematical scheme for characterizing the Schrödinger wave functions.

POSTULATE (M. Born). Let Δ denote a region of the **configuration space** Ω of a particle. Ω is the set of all coordinates $\{q\}$ where the particle can be found. For one particle, e.g., $\Omega = \mathscr{R}^3$, i.e the Euclidean space and q represents e.g. the Cartesian coordinates (x,y,z) (but one can use also spherical coordinates (r,ϑ,φ). Let $\psi(q,t)$ the Schrödinger wave function, computed by solving an appropriate, possible time dependent wave equation, describing the state of the particle. Then

$$W_\psi(\Delta) \doteq \frac{\int_\Delta |\psi(q,t)|^2 \, d\Omega_q}{\int_\Omega |\psi(q,t)|^2 \, d\Omega_q}$$

is the probability that the particle is within Δ when its coordinate q is measured ($d\Omega_q$ is an infinitesimal volume element suitable for Ω).

COMMENTS.

- An alternative interpretation of Born statistical principle reads: the particle spends a fraction $W(\Delta)$ of its lifetime in the region Δ.
- The integral

  $$\int_\Omega |\psi(q,t)|^2 \, d\Omega_q$$

  is the probability that a particle is located somewhere within Ω, and may be chosen as ”1”. This last equation represents a strong constraint on physically relevant wave functions, which must be square integrable over the configuration space, i.e. the integral of the square of $|\psi|$ must be a finite number. Physically relevant wave functions can therefore be suitably redefined (”normalized”) so that

  $$\int_\Omega |\psi(q,t)|^2 \, d\Omega_q = 1$$

- Note that, when one multiplies a wave function with a complex number with absolute value 1, one obtains a state that physically leads to the same $W_\psi(\Delta)$. Such states that differ by a phase factor are considered to be equivalent.
- The quantity

  $$W_\psi(d\Omega_q) = |\psi(q,t)|^2 \cdot d\Omega_q$$

  is the probability to find a particle in the volume $d\Omega_q$ surrounding the point q. For the Euclidean space, $d\Omega_q$ is the volume d^3x located at the apex of the position vector $\mathbf{x}$.

- $$\rho(q,t) \doteq \mid \psi(q,t) \mid^2$$
  is the probability density, i.e. a probability per unit volume to find the particle at the coordinate q.
- The – possibly complex wave function –
  $$\psi(q,t)$$
  is called the *amplitude* to find the particle at the coordinate q.
- We have presented the statistical interpretation following the original version of M. Born, having in mind the q can assume continuous values. However, there are quantum mechanical systems where q can only assume discrete values. The statistical interpretation is also valid when the configuration space is a discrete set and we will discuss this case in details later in this lecture.
- The statistical interpretation is still considered to be valid and has undergone solid experimental proof (see e.g. the Nobel prize 2022). For obtaining the discrete set of energy values of bound states of a quantum mechanical system, one can still formally proceed having still in mind Schrödinger interpretation.

4.2. Formal aspects 1: Hilbert spaces.

There is a shift of paradigm when going from classical to quantum mechanics. In classical mechanics, the information about the motion of a particle at a given energy E is stored in a graph defined over the variables (p,q) building the phase space. In quantum mechanics, instead, the "state" of a system is a string of complex numbers distributed over the coordinates q – the "wave function". We use, in the following, the symbol "q" to represents all variables $\{q_1, \ldots\ldots, q_f\}$ tha span the configuration space, i.e. the f-dimensional space that provides the f-degrees of freedom to the physical system. For instance, the configuration space of one single quantum mechanical particles in a potential $\Phi(x,y,z)$ is $\mathscr{R}^3$ $(f = 3)$. The configuration space of N point particles is the set of spatial variables $(x_1, y_1, z_1, \ldots, x_N, y_N, z_N)$ corresponding to $f = 3N$ degrees of freedom. The configuration space of a rigid body,e.g., consists of the three spatial variables identifying the center of mass and of the three angles describing its rotational degrees of freedom. Notice that a particle, like the electron, can have an internal degree of freedom (a spin) described by a variable s which can assume a discrete set of values. The mathematical structure introduced in the following covers both types of configuration spaces, those with continuous and those with discrete variables.
COMMENT: Wave functions versus "states".
The solutions of the SE are scalar functions of the variable q, denoted by $\psi(q)$. Von Neumann and Dirac describe this situation as the system being in a "state" ψ which is specified by giving a continuous string of complex values $\psi(q)$. Dirac introduces a special symbol to label states: instead of the greek letter ψ he uses the "ket" $|\psi>$. The idea behind this abstract thinking is the following: the Schrödinger representation of the physical state is one where the state is specified at the position q. However, there might be other, equivalent representations of the state, e.g. by using the variable p. The use of abstract "state" (designated by ψ or $|\psi>$) therefore offers a new degree of freedom within the formal structure of quantum mechanics.

Rule A: the superposition principle. The SE for a given solution E has the specific property that if $\psi_1(q)$ and $\psi_2(q)$ are solutions of the SE to a given value E, so is any linear combination $\alpha \cdot \psi_1(q) + \beta \cdot \psi_2(q)$. Because of this linearity, a quantum mechanical particle with a given energy eigenvalue can therefore "reside simultaneously in two or more different states" (we will specify later what this sentence exactly means). In classical mechanics, instead, a particle is prepared in some state (p_0,q_0) and then driven state-by-state – i.e. one state after the other – to other states by the time evolution provided by Newton equations, under the constraint of energy conservation. As a result one single trajectory in phase space is designed, in the specific case of a particle with hard wall the trajectory is one specific rectangle in phase space. A second trajectory (such as the one belonging e.g. to a different direction of the angular momentum vector in the classical problem of a particle moving in a radially symmetric potential) is not interfering with the first one, in the sense that when the particle is along one trajectory, it does not jump onto the other so to form a kind of mixed trajectory. Thus, while in classical mechanics the superposition of states or trajectories is meaningless – and is not even foreseen because of the non-linearity of Newton equations – in quantum mechanics the used of "mixed" states is foreseen by the linearity of the Schrödinger equation and a suitable mathematical structure must be developed that incorporates this property. This mathematical structure is the superposition principle:

SUPERPOSITION PRINCIPLE: If ψ_1 and ψ_2 are states of a system, so does the linear combination

$$\psi = \lambda_1 \cdot \psi_1 + \lambda_2 \cdot \psi_2$$

(λ_i being complex numbers).

This equation is the first rule of computation encountered when defining the fundamental mathematical structure of quantum mechanics. Because of this principle, the set of states becomes a "vector space" and the states themselves – as the elements of this vector space – are "vectors". When states are specified onto e.g. a variable q one designates $\{\psi(q)\}$ as a "function space'.

POSTULATE A: The set of states of any quantum mechanical system builds a vectors space, with the superposition principle as a computation rule.

EXAMPLES.
1. If we think of the function set

$$\{\sqrt{\frac{2}{L}}\cos\frac{(2n+1)\pi}{L}\cdot x, \sqrt{\frac{2}{L}}\sin\frac{(2n)\pi}{L}\cdot x\} \quad ; \quad n = 1,2,3,....$$

than the set of all these wave functions, including their linear combinations build a vector space of functions that have the property of vanishing at $-\frac{L}{2},\frac{L}{2}$.
2. The set

$$\{u_n(x) = \frac{1}{\sqrt{L}} \cdot e^{i\frac{2\pi\cdot n}{L}\cdot x}\} \quad ; \quad n = 0, \pm 1, \pm 2, ...$$

and all linear combinations build a vector space of periodic functions with period L.

COMMENTS.
1. The superposition principle allows to define the concept of linearly independent vectors.
DEFINITION. The vectors $\{\psi_1,...,\psi_n\}$ are linearly independent if the equation

$$\sum_1^n a_i \psi_i = 0$$

can only be satisfied by $a_1,...,a_n = 0$.
2. DEFINITION. A vector space is n-dimensional if it contains n-linearly independent vectors.
If n is finite, then the vector space is said to be finite-dimensional. If n is infinite than the vector space has infinite dimension. For instance, the space of states of one particle in a potential well is a function space. We have computed in this function space an infinite number of wave functions for the bound

states. These functions are manifestly linear independent: we conclude that the function space has infinite dimension.

Rule B: Hermitic (Hilbert) metric (scalar product). Born's postulate requires wave functions to be square integrable. A rule of computation in the vector space that takes this property into account is the "scalar product of vectors". A scalar product associates to a pair of vectors ψ, φ a scalar quantity

$$(\varphi, \psi)$$

such that

- $(\varphi, \psi) = \overline{(\varphi, \psi)}$
- $(\lambda_1 \varphi_1 + \lambda_2 \varphi_2, \psi) = \bar{\lambda}_1 (\varphi_1, \psi) + \bar{\lambda}_2 (\varphi_2, \psi)$
- $(\varphi, \lambda_1 \psi_1 + \lambda_2 \psi_2) = \lambda_1 (\varphi, \psi_1) + \lambda_2 (\varphi, \psi_2)$
- positivity: $(\varphi, \varphi) \geq 0$ and $(\varphi, \varphi) = 0 \Longleftrightarrow \varphi = 0$

In Dirac notation, the state on the left hand side is an adjunct – "bra" – vector and the scalar product writes as a "braket":

$$(\varphi, \psi) \rightarrow < \varphi \mid \psi >$$

EXAMPLE 1.
The vector space $\mathscr{C}^n$ (n being a finite integer) is a vector space with Hermitic metric. The elements $\mid \varphi >$ of this vector space are vertical columns of the type

$$\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}$$

The adjunct (transpose complex conjugate) states $< \psi \mid$ are horizontal columns of the type

$$\left(\overline{b_1}, \overline{b_2}, ..., \overline{b_n} \right)$$

and the scalar product is the matrix multiplication

$$< \psi \mid \varphi > = \left(\bar{b}_1, \bar{b}_2, ..., \bar{b}_n \right) \cdot \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} = \sum_{i=1}^{n} \bar{b}_i \cdot a_i$$

EXAMPLE 2.
In a function space one can define an Hermitic metric by

$$(\psi, \varphi) \doteq \int_{\mathscr{K}} d\Omega_q \overline{\psi(q)} \cdot \varphi(q)$$

(the case $\varphi = \psi$ (corresponding to Born postulate) being a special one). $d\Omega_q$ is the infinitesimal volume element in the configuration space and takes into account which type of variables one uses. For instance, in Cartesian coordinates we have $d\Omega_q = dx \cdot dy \cdot dz$. This integral over the configuration space involving the product of functions has all the properties that we know for the scalar product between two vectors in the vector space $\mathscr{C}^n$.

COMMENTS.
The existence of an Hermitic metric allows to define useful concepts and operations.
1. The name "metric" used in this paragraph refers to the fact that using the scalar product one can "measure" the "length" – the "norm"– of a vector ψ as

$$\| \psi \| \doteq \sqrt{(\psi, \psi)}$$

2. One can also define the distance $d(\psi,\varphi)$ between two vectors

$$d(\psi,\varphi)=\parallel\psi-\varphi\parallel$$

3. One can define orthogonal vectors as two vectors ψ, φ which fulfill

$$(\psi,\varphi)=0$$

4. A set of n-vectors $\{\psi_1,...,\psi_n\}$ with the property

$$\left(\psi_i,\psi_j\right)=\delta_{ij}$$

is called an orthonormal set (ONS).[1]

EXAMPLES.

1. The set

$$\begin{pmatrix}0\\1\\.\\.\\.\\0\end{pmatrix}\begin{pmatrix}1\\0\\.\\.\\.\\0\end{pmatrix}\quad....\quad\begin{pmatrix}0\\0\\.\\.\\.\\1\end{pmatrix}$$

builds an orthonormal set (ONS) in $\mathscr{C}^n$.

2. The set

$$\left\{\sqrt{\frac{2}{L}}\cos\frac{(2n+1)\pi}{L}\cdot q,\sqrt{\frac{2}{L}}\sin\frac{(2n)\pi}{L}\cdot q\right\}\quad;\quad n=1,2,3,....$$

provides a set with infinite orthonormal elements in the vector space of functions that vanish at $\pm\frac{L}{2}$, which is therefore an infinite dimensional vector space.

3. Similarly, in the function space consisting of L-periodic functions one can find an infinite number of mutually orthogonal functions – e.g. the set

$$\left\{u_n(x)=\frac{1}{\sqrt{L}}\cdot e^{i\frac{2\pi\cdot n}{L}\cdot x}\right\}\quad;\quad n=0,\pm1,\pm2,...$$

This function space is also infinite dimensional.

POSTULATE B: The set of states of a quantum mechanical system builds a vector space with Hermitic metric (scalar product).

Rule C: Complete orthonormal set (expansion theorem). There is a third rule of computation in quantum mechanics, called **expansion theorem**.

EXPANSION THEOREM (FINITE DIMENSIONAL VECTOR SPACE WITH SCALAR PRODUCT). Let a set of n orthonormal vectors $\{\psi_1,...,\psi_m\}$ be given.

CLAIM: any vector ψ can be expressed as a linear combination of the ψ_i:

$$\psi=\sum_{i=1}^{n}c_i\psi_i$$

---

[1] THEOREM: n-orthonormal vectors $\{\psi_1,...,\psi_n\}$ are linearly independent.

PROOF: Assume that one of the states can be expressed as a linear combination of the others:

$$\psi_j=\sum_{i=1,i\neq j}^{n}a_i\psi_i\neq0$$

Multiplying with ψ_j from the left and taking the scalar product we obtain

$$\left(\psi_j,\psi_j\right)=0$$

This leads to $\psi_j=0$ which is contrary to the assumption.

PROOF: by scalar multiplication of the expansion from the left hand side with ψ_i one obtains

$$c_i = (\psi_i, \psi)$$

c_i are the so-called "coordinates" of the state ψ. ⋄.

DEFINITION. The set $\{\psi_1, ..., \psi_m\}$ is said to build a **complete** orthonormal system of **basis states**.

COMMENTS:

Complete means that **all** basis state are known. Completeness of the orthonormal system is extremely important: only when all basis states of a quantum mechanical system are known, one has an exact and complete knowledge of all his possible quantum states. In a finite dimensional vector space, completeness is a straighforward property.

One would like to generalize the expansion theorem to an infinite dimensional vector space

$$\psi = \sum_{i=1}^{\infty} c_i \psi_i \quad ; \quad c_i = (\psi_i, \psi)$$

This is not a simple task. First, in an infinite dimensional space one has an infinite number of basis functions, and completeness is not straightforward: one must show that one has found all mutually orthogonal vectors. Second, the infinite sum must be provided with a convergence criterium.

The search for a CONS (Complete Othornormal System) in an infinite dimensional vector space and the definition of convergence are simoultaneous tasks. Assume that we have an infinite number of orthonormal states $\{\psi_i\}; i = 1, ..., \infty$ and a vector ψ. Build the partial sum

$$\sum_{i=1}^{n} c_i \psi_i \quad ; \quad c_i = (\psi_i, \psi)$$

as an approximation of the state ψ. We notice that the difference vector

$$\psi - \sum_{i=1}^{n} c_i \psi_i$$

is orthogonal to any ψ_j, as

$$\left(\psi_j, \psi - \sum_{i=1}^{n} c_i \psi_i\right) = (\psi_j, \psi) - \left(\sum_{i=1}^{n} c_i \psi_i, \psi_j\right) = (\psi_j, \psi) - \underbrace{\sum_i c_i \underbrace{(\psi_i, \psi_j)}_{\delta_{ij}}}_{c_j} = 0$$

We use this result to compute the length of the difference vector:

$$\left(\psi - \sum_{i=1}^{n} c_i \psi_i, \psi - \sum_{i=1}^{n} c_i \psi_i\right) = (\psi, \psi) - \sum_{i=1}^{n} | c_i |^2$$

(Bessel formula). Bessel formula proves the following theorem for infinite dimensional vector spaces:

THEOREM: The length of the difference vector decreases with increasing n.

DEFINITION. The set of orthonormal states $\{\psi_i\}; i = 1, ..., \infty$ is complete if, for any vector ψ, the length of the difference vector tends to zero with increasing n.

COMMENT: In general, the proof of orthonormality of a given set of states is quite easy and most used in computations. The proof of completeness, instead, might be quite "painful" and is, typically, referred to advanced textbooks of mathematical physics. For the purpose of this lecture, we recall that the physical consequences of completeness are significant. Because of completeness, the elements of a CONS build a basis set in the vector space, in the sense that their knowledge is enough to describe, by means of a linear superposition, any other element of the vector space. Accordingly, the solving of the Schrödinger equation and, in general, of quantum mechanical problems "reduces" to find a suitable CONS. Using the words of R. Feynman, physical research is about finding the "base states of the world".

POSTULATE C: The set of states $\{\psi\}$ of a quantum mechanical system is vector space with scalar product and with a complete set of orthonormal states. Such a vector space is said to be a **Hilbert**

space.
COMMENT.
Because of the relation

$$(\psi,\psi)-\lim_{n\to\infty}\sum_{i=1}^{n}|\,c_i\,|^2=0$$

and of the requirement of finiteness of (ψ,ψ) (Born), postulate C rewrites:
POSTULATE C': The set of states $\{\psi\}$ of a quantum mechanical system is vector space with scalar product and comprises all possible states with coordinates $c_1,c_2,....$ such that

$$\sum_{i=1}^{\infty}|\,c_i\,|^2$$

converges. (This is an alternative, equivalent definition of a Hilbert space).

We are now able to formulate precisely the expansion theorem in an infinite dimensional Hilbert space.
EXPANSION THEOREM (INFINITE DIMENSIONAL VECTOR SPACE WITH SCALAR PRODUCT AND CONS). Any element ψ in an Hilbert space can be written as

$$\psi=\sum_{i=1}^{\infty}(\psi_i,\psi)\psi_i$$

The "equality" means that

$$\lim_{n\to\infty}||\,\psi-\sum_{i=1}^{n}(\psi_i,\psi)\psi_i\,||^2=0$$

(normwise convergence) or, equivalently,

$$||\,\psi\,||^2=\sum_{i=1}^{\infty}|\,(\psi_i,\psi)\,|^2$$

This last expression is known as the completion relation (or Parseval identity). Its fulfillment shows that the system of basis functions used is complete, i.e. no basis function was left out. It also establishes an isometry between any Hilbert space and the Hilbert space $(c_1,c_2,....)$ with $\sum_i|\,c_i\,|^2<\infty$.
COMMENTS.
1. The use of a CONS to generate elements of the Hilbert space calls for a discussion about which functions can be reached by such an infinite sum. This is because the relation

$$\psi=\sum_{i=1}^{\infty}(\psi_i,\psi)\psi_i$$

does not necessarily imply that $\psi(q)$ and $\sum_{i=1}^{\infty}(\psi_i,\psi)\psi_i(q)$ are equal at any point q (pointwise convergence). Let, for instance, the expansion

$$\sum_i(\psi_i,\psi)\cdot\psi_i$$

be given. It is not difficult to imagine that one could alter the graph of $\psi(q)$ pointwise to produce an infinite number of singular functions that differ from the function ψ "pointwise" (more precisely: within a so called "Lebesque null set"[2]) but have the same expansion coefficients as the function ψ. Accordingly, the expansion coefficients are set into correspondence with an entire class of

---

[2] The famous Dirichelet function provides an example of such a Lebesque null set and also introduces the need to use Lebesque integration when defining functions in an Hilbert space. This function is almost everywhere zero in the interval $[0,1]$ except at the position of the rational values in the same interval, where it is given the value of "1". This function is not Riemann-integrable but it can be integrated – like all functions which have an finite or countably infinite set of point-discontinuities – by the method of Lebesque, in this case to give a vanishing integral, as the function is almost everywhere vanishing.

”equivalent” functions.[3] The elements of the Hilberts space generated by the expansion with the CONS are the ”classes of equivalent functions”.
2. According to V. Moretti (Valter Moretti, Do continuous wavefunction form a Hilbert space?, `https://physics.stackexchange.com/q/719681`), the classes contain ”true monsters”. For any practical purposes, however, (except for advanced ”proof mathematics”) one deals with representatives of the classes which are, typically continuous and at least once differentiable, so that they can be solution of the Schrödinger equation.
3. The infinite linear conbinations of continuous CONS functions might produce discontinuos functions (recall the building of, for instance, a square signal by a suitable Fourier series). Accordingly, classes with a discontinuous function as representative can also belong to the Hilbert space of a quantum mechanical system. They are, however, exotic appearances in practical physical problems related to atoms, molecules and solids (see again the comments by V. Moretti).

EXAMPLES OF CONS.
1. n orthogonal vectors in a vector space with finite dimensions n.
2. The set of infinite dimensional vectors

$$f_1 = (1,0,0,...), f_2 = (0,1,....), ...f_n = (0,0,...,1,0,0,...), ...$$

in the Hilbert space ℓ_2 of all sequences $(a_1, a_2, ...)$ for which $\sum_{i=0}^{\infty} | a_i |^2 < \infty$.
3. The set

$$\{\sqrt{\frac{2}{L}} \cos \frac{(2n+1)\pi}{L} \cdot q, \sqrt{\frac{2}{L}} \sin \frac{(2n)\pi}{L} \cdot q\} \quad ; \quad n = 1,2,3,....$$

builds a CONS in the vector space of functions that vanishes at $-\frac{L}{2}, \frac{L}{2}$.
4. A famous theorem by Fourier establishes that the plane waves

$$u_k(x) = \frac{1}{\sqrt{L}} \cdot e^{i \cdot k \cdot x}$$

with $k = \frac{2\pi}{L} \cdot n$ and $n \in \mathscr{Z}$ build a CONS in the function space of L-periodic functions.
5. The eigenfunctions

$$\left\{ \sqrt{\frac{1}{\sqrt{\pi}} \cdot \frac{1}{2^n \cdot n!}} \cdot H_n(x) \cdot e^{-\frac{x^2}{2}} \right\}$$

of the one dimensional oscillator build a CONS in the Hilbert space of square integrable functions in the variable $x \in \mathscr{R}$.

The Fourier Transform

Mathematics endows the set of square integrable functions with basis functions labeled by an index that takes continuum values: this technology is known as Fourier-transform. This is a kind of ”improper” CONS that requires a special treatement. The starting point is the Fourier expansion for L-periodic functions, that uses the basis functions

$$u_k(x) = \frac{1}{\sqrt{L}} \cdot e^{i \cdot k \cdot x}$$

---

[3] Two functions $f(q)$ and $g(q)$ belonging to the same class are considered to be equivalent $f \sim g$ if $\{q : f(q) - g(q) \neq 0\}$ is a Lebesque null set. In words: $f \sim g$ if $f(q) = g(q)$ for almost all q. This relation is an equivalence relation and therefore allows to build equivalence classes $[f]$. The set of all equivalence classes $\{[f]\}$ is a vector space with Hermitic metric:

$$[f] + [g] \doteq [f + g]; \quad \alpha[f] \doteq [\alpha \cdot f]; \quad ([f],[g]) \doteq (f,g)$$

i.e. the operations between the classes can be performed using their representatives. As each class is generated by a set of expansion coefficients, the space is complete.

with $k = \frac{2\pi}{L}\cdot n$ and $n \in \mathscr{Z}$ and orthonormality relation

$$\left(u_q, u_k\right) = \frac{1}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} dx e^{-i\cdot q\cdot x} e^{i\cdot k\cdot x} = \delta_{qk}$$

For any periodic functions one obtains the Fourier series

$$\psi(x) = \sum_k \tilde{\psi}(k)\cdot u_k(x) \quad \text{with} \quad \tilde{\psi}(k) = (u_k, \psi) = \int_{-\frac{L}{2}}^{\frac{L}{2}} dx\, \overline{u_k(x)}\cdot\psi(x)$$

We would like to find out how the Fourier series develop in the limit $L \to \infty$, so that it covers the space of physical functions defined over the entire configuration space $\mathscr{R}$, also without imposing periodic boundary conditions. For this purpose we write

$$\psi(x) = \sum_k \tilde{\psi}(k) u_k(x) = \frac{1}{\sqrt{2\pi}}\sum_k \frac{\sqrt{L}}{\sqrt{2\pi}}\cdot\tilde{\psi}(k)\cdot\frac{2\pi}{L}\cdot e^{ikx} = \frac{1}{\sqrt{2\pi}}\sum_k \tilde{\psi}'(k)\Delta_k e^{ikx}$$

with

$$\tilde{\psi}'(k) = \frac{1}{\sqrt{2\pi}}\int_{-\frac{L}{2}}^{\frac{L}{2}} dx\cdot\psi(x)\cdot e^{-ikx} \quad \text{and} \quad \Delta_k = \frac{2\pi}{L}$$

Let now $L \to \infty$: then $\Delta_k \to 0$ and the Riemann sum

$$\frac{1}{\sqrt{2\pi}}\sum_k \tilde{\psi}'(k)\Delta_k e^{ikx}$$

converges to

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\cdot\tilde{\psi}(k)\cdot e^{ikx}$$

with

$$\tilde{\psi}(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\cdot\psi(x)\cdot e^{-ikx}$$

This equation tells us e.g. that any square integrable function $\psi(x)$ can be expressed as a **Fourier-integral**. The square integrable function $\tilde{\psi}(k)$ is its Fourier transform. The Parseval identity writes

$$\int_{-\infty}^{\infty} dx \mid\psi(x)\mid^2 = \int_{-\infty}^{\infty} dk \mid\tilde{\psi}(k)\mid^2$$

and is called the Plancherel relation.

Dirac delta function technology. The Fourier transform technology introduced in the last subsection is an example of an expansion of a function into a set of basis functions labeled by a continuous index, namely k. In fact, one can interpret the expression

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\cdot\psi(k)\cdot e^{ikx}$$

formally as the superposition of basis states u_k with

$$u_k(x) = \frac{1}{\sqrt{2\pi}}\cdot e^{ik\cdot x} \quad ; \quad k \in \mathscr{R}$$

The basis states themselves are not square integrable and cannot be realized as physical states. However, we can use them provided they gives physically plausible results. In order to use them as basis states, we need to establish the meaning of the orthonormality relation

$$(u_{k'}, u_k) = \frac{1}{2\pi}\int_{\mathscr{R}} dx\, e^{-ik'x}\cdot e^{ikx}$$

This is a very important integral that must be discussed with some care. The result of this integral, in fact, must be an expression containing k, k' (by integration, the variable x is eliminated). But the expression we will obtain **is not a function** of k, k' in the usual sense, but a "generalized function". Such a generalized function (or distribution) was first introduced by Dirac and later a very rigorous mathematical framework was worked out by L. Schwartz and I.M. Gel'fand. We will learn how to use it without being "too" rigorous. Let us now compute the integral as

$$\lim_{L\to\infty}\left[\frac{1}{2\pi}\cdot\int_{-L}^{L}dx\cdot e^{ix\cdot(k-k')}\right]$$

First we use the Euler formula to write

$$e^{ix\cdot(k-k')}=\cos(x\cdot(k-k'))+i\cdot\sin(x\cdot(k-k'))$$

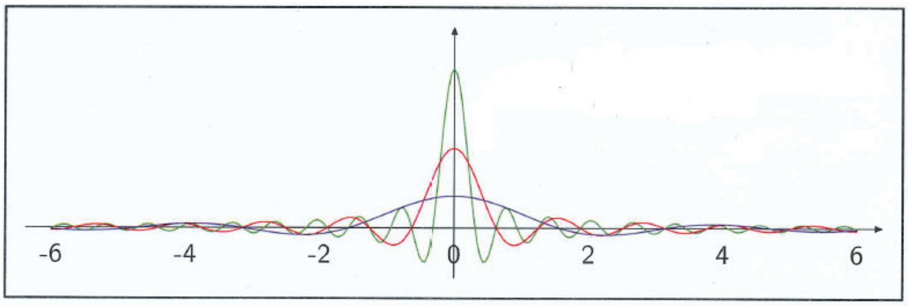
We point out that the sin-function is uneven and its integral from $-L$ to $+L$ vanishes for symmetry reasons. We are left with the elementary integral over the cos-functions. This gives:

$$\lim_{L\to\infty}\frac{1}{2\pi}\int_{-L}^{L}\cos(x\cdot(k-k'))dx=\lim_{L\to\infty}\frac{1}{\pi}\frac{\sin(L\cdot(k-k'))}{(k-k')}$$

The function

$$\frac{1}{\pi}\frac{\sin(L\cdot(k-k'))}{(k-k')}$$

has some interesting properties (see the graph):



The function $\frac{1}{\pi}\frac{\sin(L\cdot(k-k'))}{(k-k')}$ *plotted as a function of* $(k-k')$ *for* $L=2$ *(black line),* $L=5$ *(red line) and* $L=20$ *(green line).*

- for $k=k'$ it amounts to $\frac{L}{\pi}$. Accordingly, for $L\to\infty$ it diverges at $k=k'$.
- For $k\neq k'$ it oscillates with the period $\frac{2\pi}{L}$. Accordingly, for $L\to\infty$ it "averages" to zero for $k\neq k'$.
- Furthermore

  $$\int_{-\infty}^{\infty}dk\frac{1}{\pi}\frac{\sin((k-k')\cdot L)}{(k-k')}=\int_{-\infty}^{\infty}dk'\frac{1}{\pi}\frac{\sin((k-k')\cdot L)}{(k-k')}=1$$

  for any L.

DEFINITION: The function that
1. diverges when its argument is 0;
2. vanishes elsewhere so that
3. its integral is"1"
is the so called Dirac Delta function, $\delta(x)$.

It is used, among other, to write the orthonormality relation for plane waves:

$$(u_{k'},u_k)=\frac{1}{2\pi}\int_{\mathscr{R}} dx\, e^{-ik'x}\cdot e^{ikx}\doteq\delta(k-k')$$

$$\delta(k-k')=\begin{cases}\infty, \quad k=k'\\ 0, \quad k\neq k'\\ \int dk\cdot\delta(k-k')=1\end{cases}$$

THEOREM. For any continuos, integrable function $f(x)$ we have

$$\int_{\infty}^{\infty} dx\cdot\delta(x)\cdot f(x)=f(0)$$

and in general

$$\int_{\infty}^{\infty} dx\cdot\delta(x-y)\cdot f(x)=f(y)$$

and

$$\int_{\infty}^{\infty} dy\cdot\delta(x-y)\cdot f(y)=f(x)$$

i.e. the delta-function is actually a functional that associates to every continuous function $f(x)$ its value $f(0)$ (this result points out that the Dirac function is a so-called ”distribution” or ”generalized function”: a ”distribution” is also a ”functional”.)

”PROOF”: Approximate e.g. $f(x)$ in the vicinity of $x=0$ with $f(0)+O(x)$ and write

$$\begin{aligned}\int_{-\infty}^{\infty} dx\delta(x)\cdot f(x) &= \lim_{k\to\infty}\int_{-\infty}^{\infty} dx\frac{1}{\pi}\frac{\sin k\cdot(x)}{(x)}\cdot f(x)\\ &= f(0)\cdot\underbrace{\lim_{k\to\infty}\int_{-\infty}^{\infty} dx\frac{1}{\pi}\frac{\sin k\cdot(x)}{(x)}}_{1}+\lim_{k\to\infty}\int_{-\infty}^{\infty} dx\frac{1}{\pi}\frac{\sin k\cdot(x)}{(x)}\cdot O(x)\\ &= f(0)+O\left(\lim_{k\to\infty}\int_{-\infty}^{\infty} dx\frac{1}{\pi}\cdot\sin(k\cdot x)\right)\end{aligned}$$

We compute now the rest integral:

$$\lim_{k\to\infty}\int_{-\infty}^{\infty} dx\frac{1}{\pi}\cdot\sin(k\cdot x)=\frac{1}{\pi}\lim_{k\to\infty}\left(-\frac{\cos(k\cdot L)}{k}\Big|_{L\to-\infty}^{L\to\infty}\right)$$

The cos-function is limited to the ±1 interval, so that the limit $k\to\infty$ produces the vanishing of the rest integral. We are now left with the sought for result.

This is the summary of useful formulas:

$$\boxed{\delta(x)=\begin{cases}\infty\,;\, x=0\\ 0\,;\, x\neq 0\\ \int_{-\infty}^{\infty} dx\cdot\delta(x)=1\end{cases}} \qquad \boxed{\int_{\infty}^{\infty} dx\cdot\delta(x)\cdot f(x)=f(0)}$$

$$\boxed{(u_{k'},u_k)=\frac{1}{2\pi}\int_{\mathscr{R}} dx\, e^{-ik'x}\cdot e^{ikx}\doteq\delta(k-k')}$$

$$\boxed{\psi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\cdot\tilde{\psi}(k)\cdot e^{ikx}} \qquad \boxed{\tilde{\psi}(k)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\cdot\psi(x)\cdot e^{-ikx}}$$

4.3. Formal aspect 2: Operators

4.3.1. Introduction

By virtue of the correspondence principle, there are further mathematical objects (besides wave functions) that play a central role in quantum mechanics: the so called ”operators”. This principle is, however, an heuristic one. In the following, we obtain operators by starting from a more fundamental principle, the statistical interpretation of the wave function by M. Born.

Once a probability amplitude $\psi(\mathbf{x})$ is defined, one can ask about the result of performing one single act of determining the particle position. The answer is that the particle will be registered at the site $\mathbf{x}_1$ with the probability $dV_1 \mid \psi(\mathbf{x}_1) \mid^2$. Making sure that the particle is prepared into the same state ψ, one can repeat the measurement with a detector placed at the site $\mathbf{x}_2$. The particle will be registered at the site $\mathbf{x}_2$ with the probability $dV_2 \cdot \mid \psi(\mathbf{x}_2) \mid^2$, and so on. Upon repeating the measurement many times the outcome of the position determination will approach asymptotically the ”average” position[4]

$$\bar{\mathbf{x}}\,|_{\psi} \doteq \int_{\mathscr{R}} dV \cdot \mathbf{x} \cdot \mid \psi(\mathbf{x}) \mid^2$$

In the framework of quantum mechanics, one can obtain the same result for the average position resorting to introducing an operator $\hat{\mathbf{q}}$ that maps any state ψ into a new state $\hat{\mathbf{q}}\psi$, defined by evaluating the result at any point $\mathbf{x}$:

$$(\hat{\mathbf{q}}\psi)(\mathbf{x}) = \mathbf{x} \cdot \psi(\mathbf{x})$$

The map is linear, i.e. it preserves the linearity of the vector space:

$$\hat{\mathbf{q}}(\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1\hat{\mathbf{q}}\psi_1 + \lambda_2\hat{\mathbf{q}}\psi_2$$

(this equation can be proved by evaluating its validity at any point $\mathbf{x}$). Linearity of operators is essential in order to keep the validity of the superposition principle. Once such an operator is defined, one can compute the scalar product of $\hat{\mathbf{q}}\psi$ onto ψ:

$$(\psi, \hat{\mathbf{q}}\psi) = \int_{\mathscr{R}} dV \cdot \psi^*(\mathbf{x}) \cdot \mathbf{x} \cdot \psi(q) \doteq < \mathbf{q} >_{\psi}$$

This scalar product is called the quantum mechanical expectation value of the operator $\hat{\mathbf{q}}$ in the state ψ. By explicit computation, we have just established the identity

$$\bar{\mathbf{x}}_{\psi} = < \hat{\mathbf{q}} > |_{\psi}$$

We have therefore succeeded in expressing the statistical average by means of the expectation value of an operator. It appears therefore that a classical physical quantity such as the position vector has a counterpart in quantum mechanics, in agreement with the heuristic correspondence principle.

To corroborate this hypothesis we would like to find such an operator for the momentum of the particle as well. In order to keep the mathematics as comprehensible as possible we consider a single particle moving along the x-axis. Let the state of the particle being given by ψ. We are used to specifying this state at any value of x but we recall that Dirac and v. Neumann has opened the possibility of specifying this state over the variable p. $\psi(p)$ would then be the amplitude that the momentum of the particle assumes the value p. Accordingly, we can write

$$\bar{p}\,|_{\psi} = \int_{-\infty}^{\infty} dp\, \overline{\psi(p)} \cdot p \cdot \psi(p)$$

---

[4] In this expression we have assumed that the state ψ has been normalized to 1, i.e. $\int_{\mathscr{R}} dV \cdot \mid \psi(\mathbf{x}) \mid^2 = 1$, otherwise the statistical average amounts to

$$\bar{\mathbf{x}}\,|_{\psi} = \frac{\int_{\mathscr{R}} dV \cdot q \cdot \mid \psi(\mathbf{x}) \mid^2}{\int_{\mathscr{R}} dV \cdot \mid \psi(\mathbf{x}) \mid^2}$$

For the sake of simplicity we will continue with normalized states.

We would like to compute this statistical average as the quantum mechanical expectation value of a suitable operator $\hat{p}$ over the state $\psi(x)$, i.e. to write

$$\int_{-\infty}^{\infty} dp\,\overline{\psi(p)}\cdot p\cdot\psi(p) \quad \text{as} \quad (\psi,\hat{p}\,\psi)=\int_{\mathscr{R}} dx\,\overline{\psi(x)}(\hat{p}\psi)(x)$$

To solve this task we use

$$\psi(p)=\frac{1}{\sqrt{2\pi\cdot\hbar}}\int_{-\infty}^{\infty} dx\cdot\psi(x)\cdot e^{-i\frac{p}{\hbar}x}$$

and

$$\begin{aligned}\int_{-\infty}^{\infty} dp\,\overline{\psi(p)}\cdot p\cdot\psi(p) &= \int_{-\infty}^{\infty} dp\left[\frac{1}{\sqrt{2\pi\cdot\hbar}}\int_{-\infty}^{\infty} dx\cdot\overline{\psi(x)}\cdot e^{i\frac{p}{\hbar}x}\right]\cdot p\cdot\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dy\cdot\psi(y)\cdot e^{-i\frac{p}{\hbar}y}\right]\\ &= \frac{1}{2\pi\cdot\hbar}\int_{-\infty}^{\infty} dx\cdot\overline{\psi(x)}\cdot\left[\left[\int_{-\infty}^{\infty} dy\cdot\int_{-\infty}^{\infty} dp\cdot e^{i\frac{p}{\hbar}x}\cdot p\cdot e^{-i\frac{p}{\hbar}y}\cdot\psi(y)\right]\right.\end{aligned}$$

We evaluate

$$\int_{-\infty}^{\infty} dy\int_{-\infty}^{\infty} dp\cdot e^{i\frac{p}{\hbar}x}\cdot p\cdot e^{-i\frac{p}{\hbar}y}\cdot\psi(y)$$

using

$$p\cdot e^{-i\frac{p}{\hbar}y}=i\cdot\hbar\cdot\frac{\partial}{\partial y}e^{-i\frac{p}{\hbar}y}$$

and partial integration:

$$\int_{-\infty}^{\infty} dy\cdot i\cdot\hbar\cdot\frac{\partial}{\partial y}e^{-i\frac{p}{\hbar}y}\cdot\psi(y)=e^{-i\frac{p}{\hbar}y}\cdot\psi(y)\,|_{-\infty}^{\infty}+\int_{-\infty}^{\infty} dy\cdot e^{-i\frac{p}{\hbar}y}\left[-i\cdot\hbar\cdot\frac{\partial}{\partial y}\right]\psi(y)$$

The first term of the right hand side of this last equation vanishes because of the square integrability of the amplitude $\psi(y)$, leaving

$$\begin{aligned}\int_{-\infty}^{\infty} dp\,\overline{\psi(p)}\cdot p\cdot\psi(p) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx\cdot dy\cdot\psi^{*}(x)\cdot\left[-i\cdot\hbar\cdot\frac{\partial}{\partial y}\right]\psi(y)\cdot\underbrace{\frac{1}{2\pi\cdot\hbar}\cdot\int_{-\infty}^{\infty} dp\cdot e^{i\frac{p}{\hbar}(x-y)}}_{\frac{1}{2\pi}\cdot\int_{-\infty}^{\infty} dk\cdot e^{ik(x-y)}=\delta(x-y)}\\ &= \int_{-\infty}^{\infty} dx\cdot\overline{\psi(x)}\cdot\left[-i\cdot\hbar\cdot\frac{\partial}{\partial x}\right]\psi(x)\end{aligned}$$

We recognize that we can express $\bar{p}\,|_{\psi}$ as $<\psi,\hat{p}\,\psi>$ with

$$(\hat{p}\,\psi)(x)\doteq -i\cdot\hbar\cdot\frac{\partial}{\partial x}\psi(x)$$

This result is in agreement with the heuristic correspondence principle.

EXAMPLES OF OPERATORS.

1. The position operator is defined as

$$\hat{\mathbf{q}}\psi(\mathbf{x})=\mathbf{x}\cdot\psi(\mathbf{x})$$

2. The momentum operator is defined as

$$\hat{\mathbf{p}}\psi(\mathbf{x})=-i\cdot\hbar\nabla\psi(\mathbf{x})$$

3. The operator of the orbital angular momentum is defined as

$$\hat{\mathbf{L}}=\hat{\mathbf{x}}\times\hat{\mathbf{p}}$$

This is a vector operator with Cartesian components

$$(\hat{L}_x)\psi(\mathbf{x})=-i\hbar\left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right)\psi(\mathbf{x})$$

$$(\hat{L}_y)\psi(\mathbf{x}) = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\psi(\mathbf{x})$$

$$(\hat{L}_z)\psi(\mathbf{x}) = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi(\mathbf{x})$$

4. For a particle with mass m in a potential field $\Phi(\mathbf{x})$ the classical Hamilton function – which expresses the total energy in the classical variables $(\mathbf{x},\mathbf{p})$ reads

$$H(\mathbf{x},\mathbf{p}) = \frac{\mathbf{p}^2}{2m} + \Phi(\mathbf{x})$$

One can construct an **Hamilton operator** by building $\hat{H}(\hat{\mathbf{x}},\hat{\mathbf{p}})$:

$$\left(\hat{H}(\hat{\mathbf{x}},\hat{\mathbf{p}})\right)\psi(\mathbf{x}) = -\frac{\hbar^2}{2m}\triangle\psi(\mathbf{x}) + \Phi(\mathbf{x})\cdot\psi(\mathbf{x})$$

5. Functions of $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$. Given a function $f(\mathbf{x})$ one can define the operator $\hat{f}(\hat{\mathbf{x}})$ by

$$(\hat{f}\psi(\mathbf{x}) \doteq f(\mathbf{x})\cdot\psi(\mathbf{x})$$

Similarly, one can define formally polynomial operators (or any function that can be written as power series) of $\hat{\mathbf{p}}$ by means of

$$(\hat{f}\psi(\mathbf{x}) \doteq f(-i\cdot\hbar\frac{\partial}{\partial\mathbf{x}})\cdot\psi(\mathbf{x})$$

Operator calculus In general, a linear operator is a "device" that maps a state ψ into $\hat{A}\psi$, such that

$$\hat{A}(a_1\psi_1 + a_2\psi_2) = a_1\hat{A}\psi_1 + a_2\hat{A}\psi_2$$

An operator is specified by evaluating

$$(\hat{A}\psi)(\mathbf{x})$$

$\hat{A}$ is typically some differential or integral operator.

The set of linear operators builds an algebra, i.e. operators can be composed with the two operations of "sum" and "product", with the following computation rules:

$$\begin{aligned}(\hat{A}+\hat{B})\psi &= \hat{A}\psi + \hat{B}\psi\\ (\hat{A}\cdot\hat{B})\psi &= \hat{A}(\hat{B}\psi)\\ (\lambda\cdot\hat{A})\psi &= \hat{A}(\lambda\cdot\psi)\end{aligned}$$

An operator is only well defined if its domain of definition $D(\hat{A})$ is also specified. The domain of definition establishes the set of states onto which the operator can be applied. It is related to some boundary conditions. A consequence of attaching boundary conditions to some given operator is that two operators might be formally identical but have a different domain of definition. For instance, the operator $-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$ applied to functions that vanishes at $\pm\frac{L}{2}$ has different eigenvalues than the formally identical operator with periodic boundary conditions. Notice that $D(\hat{A})$ needs not to be the entire Hilbert space, but if φ is any state in the Hilbert space, there is a state $\psi \in D(\hat{A})$ that is arbitrarily close to φ. For instance, the momentum operator $-i\frac{d}{dx}$ in one dimension is only defined for square integrable functions which are also differentiable with respect to the variable x, while the Hilbert space might contain functions that are not everywhere differentiable.

DEFINITION.

The quantum mechanical expectation value of the operator $\hat{A}$ in a given state ψ writes:

$$<\hat{A}>|_\psi \doteq \frac{\left(\psi,\hat{A}\psi\right)}{(\psi,\psi)} \in \mathscr{C}$$

In Dirac notation:

$$<\hat{A}>|_{|\psi>} = \frac{<\psi\,|\hat{A}|\,\psi>}{<\psi\,|\,\psi>}$$

4.3.2. Matrix representation of operators (W. Heisenberg)

In virtue of the expansion theorem one can achieve a matrix representation of linear operators . Take $f = \sum_{j=1}^{\infty} a_j f_j$ and let the operator $\hat{A}$ map this state into the state $g = \sum_{j=1}^{\infty} b_j f_j$. Our aim is to write the equation

$$\hat{A} f = g$$

as a relation involving the coordinates $\{a_i\}$ and $\{b_i\}$. For this purpose we insert the expansions:

$$\hat{A} \sum_{j=1}^{\infty} a_j f_j = \sum_{j=1}^{\infty} b_j f_j$$

We multiply both sides from the left with f_i and take the scalar product, obtaining

$$\underbrace{\left(f_i, \sum_{j=1}^{\infty} a_j \hat{A} f_j \right)}_{\sum_{j=1}^{\infty} (f_i, \hat{A} f_j) \cdot a_j} = \underbrace{\sum_{j=1}^{\infty} b_j \underbrace{(f_i, f_j)}_{\delta_{ij}}}_{b_i}$$

The left-hand side contains new quantities: the so called matrix elements of the operator $\hat{A}$ with respect to the base states $\{f_i\}$:

$$A_{ij} = \left(f_i, \hat{A} f_j \right)$$

and the sought for relation between $\{a_i\}$ and $\{b_i\}$ reads

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} & .. \\ A_{21} & A_{22} & \cdots & A_{2n} & .. \\ .. & .. & \cdots & \cdots & \cdots \\ .. & .. & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} & .. \\ .. & .. & \cdots & \cdots & .. \\ .. & .. & \cdots & \cdots & .. \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_n \\ . \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \\ . \end{pmatrix}$$

DEFINITION. The matrix

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} & .. \\ A_{21} & A_{22} & \cdots & A_{2n} & .. \\ .. & .. & \cdots & \cdots & \cdots \\ .. & .. & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} & .. \\ .. & .. & \cdots & \cdots & .. \\ .. & .. & \cdots & \cdots & .. \end{pmatrix}$$

is the matrix representation of the operator $\hat{A}$ with respect to the base states $\{f_i\}$. In the following we will use the same symbol $\hat{A}$ to label both the operator and its matrix representation.

4.4. Observables

In the context of classical physics several quantities appear that are the subject of measurement intended to detect their value. Referring to one single particle for simplicity, measurable quantities are e.g. the position of a particle, its momentum, its energy and its angular momentum. The measurement instruments consists e.g. of cloud chambers for detecting the position of a particle, calorimeters for measuring the energy of particles, speed sensor for detecting the velocity, polarizers and analyzer for detecting the polarization state, etc. For such measurable quantities we adopt the term of ”observables”, which was used e.g. by Dirac in his attempt to find such measurable quantites within the realm of wave mechanics (Landau and Lifshitz use the terminology ”physical quantity”).

Our experience with the correspondence principle tells us that observables will be a subset of the linear operators. The type of experiment we have ideally performed to define the position operator is one where an act of detection is performed and the outcome is a probability density $|\psi(q)|^2$ to find the particle in q. If we restore the original state again and again after having performed again a single act of detection, the outcome of the experiment will approach asymptotically the quantum mechanical expectation value of the position operator operator in the state ψ, in line with Born postulate. This outcome sets the accent onto the quantum mechanical expectation value as a way of gauging whether an operator corresponds to an observable or not. In fact, an obvious necessary condition for operators to represent observables is that their quantum mechanical expectation value, taken over any of the elements onto which the operator can be applied, is a real number. The requirement that the expectation value is a real number reads

$$\left(\psi,\hat{A}\psi\right)=\overline{\left(\psi,\hat{A}\psi\right)}$$

Using the property of scalar product, the right hand side can be transformed to

$$\left(\hat{A}\psi,\psi\right)$$

so that the requirement that the expectation value is a real number can be written as an equation that the operator $\hat{A}$ must fulfill, for any ψ to which A can be applied (i.e. in the domain of $\hat{A}$):

$$\left(\psi,\hat{A}\psi\right)=\left(\hat{A}\psi,\psi\right)$$

The set of linear operators for which this last equation hold are called symmetric operators. We have just proven following

Theorem: A necessary condition for operators to represent observables is that they are symmetric.

EXAMPLES OF SYMMETRIC OPERATORS.

1. The operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ are symmetric. The proof of this is given by partial integration. The boundary integrals occurring upon partial integration vanishes if the wave functions are required to vanish at infinity. The vanishing of the wave functions at infinity is necessary in order to ensure that the norm of the functions is finite.
2. The Hamilton operator $\hat{H}(\mathbf{x},\mathbf{p})$ for atoms and molecules is symmetric (T. Kato, Trans. Amer. Math. Soc. **70**, p.195 (1951)).
3. A further example of symmetric operator is e.g. the Hilbert-Schmidt integral operator

$$\hat{F}\psi(x)\doteq\int_a^b dyG(x,y)\psi(y)$$

provided the kernel obeys the relation $G(x,y)=\overline{G(y,x)}$.

Proof:

$$\langle\varphi(x),F\psi(x)\rangle=\int_a^b\int_a^b dxdy\overline{\varphi(x)}G(x,y)\psi(y)$$

$$=\int_a^b\int_a^b dxdy\overline{[G(y,x)\varphi(x)]}\psi(y)=\int_a^b\int_a^b dxdy\overline{[G(x,y)\varphi(y)]}\psi(x)=\langle F\varphi(x),\psi(x)\rangle$$

Whether all symmetric operators are also ”observables” depends on more subtle details of the relation between quantum mechanics and the outcome of an experiment. We first discuss these details for the simple case of a finite dimensional Hilbert space. The infinite dimensional case is, mathematically speaking, much more difficult to handle an we will only quote the main results.

4.4.1. The spectral theorem and the outcome of a measurement (finite dimensional space).

Let the space have a dimension n and let a CONS $\{e_i\}$ be defined in it. The CONS are a set of wave functions that fulfills some boundary condition. If the operator is symmetric, its matrix elements have the property

$$\overline{A_{ji}} = \overline{(e_j, Ae_i)} = \overline{(Ae_j, e_i)} = (e_i, Ae_j) = A_{ij}$$

i.e. symmetric operators are represented by so called Hermitic matrices. These are matrices for which the transpose complex conjugate is identical with the matrix itself.
THEOREM 1: The eigenvalues of a symmetric matrix are real.
THEOREM 2: Eigenfunctions of a symmetric matrix belonging to different eigenvalues are orthogonal.

PROOF OF 1: Let e be an eigenfunction of $\hat{A}$, i.e. $\hat{A}e = \lambda e$. Than, from

$$(e, \hat{A}e) = \lambda(e,e) = (\hat{A}e, e) = \overline{\lambda}(e,e)$$

we obtain $\lambda = \overline{\lambda}$.
PROOF OF 2: Let e_1 and e_2 be eigenfunctions of $\hat{A}$ belonging to different eigenvalues λ_1, λ_2: $\hat{A}e_{1,2} = \lambda_{1,2}e_{1,2}$. By the requirement of symmetricity we have

$$(e_1, \hat{A}e_2) = (\hat{A}e_1, e_2) \Longleftrightarrow (\lambda_1 - \lambda_2)(e_1, e_2) = 0 \Longleftrightarrow (e_1, e_2) = 0$$

SPECTRAL THEOREM (finite dimensional Hilbert space): An $n \times n$ Hermitic (or symmetric) matrix has n-real eigenvalues and n-corresponding, mutually orthogonal eigenvectors, i.e. it can be brought into diagonal form, all diagonal matrix elements (the eigenvalues of the matrix) being real.
PROOF.
The proof of the spectral theorem is typically performed by induction. First we show that the theorem holds for $n = 1$. In $n = 1$ $\hat{A}$ is the multiplication by a scalar $\hat{A}a = \lambda a\ \forall a$, λ being a real number. The eigenvalue is λ and the number 1 is the eigenvector. Thus, the theorem is true in this particular case. We now assume that the theorem holds for dimension $n-1$. Because of the fundamental theorem of algebra, which asserts that the determinantal equation of $\hat{A}$ has at least one non-zero solution, we know that at least one non-zero eigenvalue λ_n exists, and we call the corresponding eigenfunction f_n. We now perform the Gram Schmidt orthogonalization procedure to construct a set of vectors $f_1, ..., f_{n-1}$ orthogonal to f_n. This means that the matrix elements

$$\hat{A}_{nj} \doteq (f_n, \hat{A}f_j) = (\hat{A}f_n, f_j) = \lambda_n (f_n, f_j) = 0 = A^*_{jn}$$

are vanishing and in the orthonormal basis $f_1, f_2, ...f_n$ the matrix representation of $\hat{A}$ reads

$$\begin{pmatrix} [A_{n-1}] & [0] \\ [0] & \lambda_n \end{pmatrix}$$

[] indicates $(n-1) \times (n-1)$ block matrices. As we assumed that the $[A_{n-1}]$ matrix can be diagonalized, we have proven the spectral theorem for finite Hermitic matrices.

COMMENTS:
1. The space belonging to one single eigenvalue might be one-dimensional or multi-dimensional. The number n_λ of eigenvectors to the eigenvalue λ is called the degeneracy of the eigenvalue. Given n_λ linearly independent vectors within the eigenspace of λ, the Gram–Schmidt ortho-normalization algorithm allows to construct an orthonormal set within each finite eigenspace itself. The entire set of orthonormalized eigenvectors builds a CONS.
2. In virtue of this theorem, in the basis set in which $\hat{A}$ is diagonal, we have, for any e

$$e = \sum_{k=1}^{m} \sum_{n=1}^{n_k} \underbrace{(f_{k,n}, e)}_{a_{kn}} f_{k,n}$$

n_k being the degeneracy of the k-th eigenvalue and the completeness relation reads

$$\| e \| = \sqrt{\sum_{k,n} | a_{kn} |^2}$$

This relation expresses the fact that the set $\{a_{kn}\}$ is enough to express the length of any vector, i.e. no coordinate (no basis vector) was left out.

Outcome of a single act of measurement. The main consequence of the existence of a spectral decomposition of the Hermitic matrix $\hat{A}$ is the outcome of a single act of measurement performed in a state e. In fact, the spectral decomposition can be used to compute the quantum mechanical expectation value of $\hat{A}$ in the state e. Write

$$e = \sum_{k=1}^{m} \sum_{n=1}^{n_k} a_{kn} f_{k,n}$$

and compute the result of applying the operator $\hat{A}$ onto the state e:

$$\hat{A}e = \sum_{k=1}^{m} \sum_{n=1}^{n_k} \lambda_k \cdot a_{kn} f_{k,n}$$

Use this equation to compute the expectation value of $\hat{A}$ in the state e:

$$(e,\hat{A}e) = \sum_{k=1}^{m} \sum_{n=1}^{n_k} \sum_{k'=1}^{..} \sum_{n'=1}^{n_{k'}} \overline{a_{k'n'}} \cdot a_{kn} \cdot \lambda_k \cdot \underbrace{(f_{k'n'}, f_{k,n})} \delta_{kk'} \cdot \delta_{nn'} = \sum_{k=1}^{m} \sum_{n=1}^{n_k} \lambda_k \cdot | a_{kn} |^2$$

COMMENTS.

1. The right hand side indicates that the building of the expectation value takes place by a set of single acts of measurements, the outcome of which is one of the eigenvalues λ_k. The probability of measuring λ_k in a single act of measurement is

$$\sum_{n=1}^{n_k} | (f_{kn}, e) |^2$$

In other words: the set of eigenvalues of the symmetric matrix $\hat{A}$ contains all possible outcomes of a measurement for the associated observable A.

2. For example, if the state e is an eigenstate, then the result of the measurement is the eigenvalue itself with a probability of 1.

3. Through this result, the coefficients forthcoming in the expansion of e in terms of the eigenstates of the operator $\hat{A}$ have a central significance for the outcome of a measurement. Roughly speaking: a_{kn} are the amplitudes that a system in the state e possesses the "typical features" of the eigenstate f_{kn} of the observable A under consideration.

4. This result establishes that the eigenvalue problem of an operator and the set of eigenvalues are playing a central role in quantum mechanics.

4.4.2. HILBERT-v. NEUMANN SPECTRAL THEOREM for infinite dimensional Hilbert spaces (without proof).

On the base of the previous result for finite dimensional spaces we update our definition of an observable as follows:

DEFINITION: An operator represents an observable if and only if it has a complete set of eigenvalues and eigenfunctions, i.e. if the spectral decomposition is complete.

This definition can be certainly adopted in infinite dimensional Hilbert spaces, where most of the quantum mechanical systems actually reside. The question is, however, about which class of operators has a complete set of eigenvalues and eigenvectors. The spectral theorem foresees that, for finite dimensional Hilbert spaces, symmetric matrices are the sought for class of operators representing observables. However, the proof of the spectral theorem for the finite dimensional case relied on the existence of at least one eigenvalue. The one eigenvalue was found from the main theorem of algebra, which cannot be applied to polynomials with an infinite power. Accordingly, the generalization of the spectral theorem to infinite dimensional Hilbert spaces is not straightforward.

Let us make some examples of the difficulties one encounters when working with symmetric operators in infinite dimensional function spaces. For the spectrum (i.e. the set of eigenvalues) of a symmetric operator in an infinite dimensional Hilbert space we can have, in fact, following situations:

•**The spectrum is an empty set.** Consider e.g. the differential operator

$$-i\frac{d}{dx}$$

and define it within the function space of differentiable square integrable functions $f(x)$ in the interval $[0,L]$. The operator is not symmetric:

$$\begin{aligned}(f,\hat{p}g) &= \int_0^L dx\bar{f}(x)\cdot[-i\frac{d}{dx}g(x)] \overbrace{=}^{\text{P.}I.} -i\bar{f}(x)\cdot g(x)\,|_0^L + \int_0^L dx\overline{(-i\frac{d}{dx}f(x))}\cdot g(x) \\ &= -i\bar{f}(x)\cdot g(x)\,|_0^L + (\hat{p}f,g) \neq (\hat{p}f,g)\end{aligned}$$

as the constant $-i\overline{f(x)}\cdot g(x)\,|_0^L$ arising from the integration by parts is, in general not vanishing. However, the operator can be rendered symmetric by imposing some boundary conditions. A possible set of boundary conditions is to allow only functions that vanish at $x=0$ and $x=L$, so that the constant $-i\overline{f(x)}\cdot g(x)\,|_0^L$ arising from the integration by parts is made to be vanishing. Under this boundary condition, however, the eigenvalue equation

$$-i\frac{d}{dx}f(x)=\lambda f(x)$$

has no solution, as the only functions solving the differential equation and fulfilling the boundary condition is the function $f(x)\equiv 0$, which cannot be considered as an eigenfunction: the operator, with the boundary condition introduced, has neither eigenfunctions nor eigenvalues. According to our understanding of the measurement process, it cannot be an observable, although it is symmetric.

• **Infinite number of countable eigenvalues.** This is e.g. the case of the Hamilton operator of a particle in a one-dimensional motion with boundary conditions $\psi(-\frac{L}{2})=\psi(\frac{L}{2})=0$ or periodic boundary conditions. It is also the case of the quantum mechanical harmonic oscillator. Consider further the operator $-i\frac{d}{dx}$. With periodic boundary conditions it is a symmetric operator. Enforcing the specific boundary condition of L-periodicity onto the solution $e^{i\lambda x}$ of the eigenvalue equation $-i\frac{d}{dx}f(x)=\lambda f(x)$ produces an algebraic equation $e^{i\lambda L}=1$ for the sought for eigenvalues λ which is solved by the infinite set of discrete eigenvalues $\lambda_n\frac{2\pi}{L}\cdot n\ n\in\mathscr{Z}$ and the infinite set of orthonormalized eigenfunctions $\frac{1}{\sqrt{L}}e^{i\lambda_n\cdot x}$.

In the situation that the infinite, countable set of eigenfunctions is complete – i.e. that there are no further functions orthogonal to the discrete spectrum eigenfunctions – the spectral theorem for finite dimensional can be generalized to this case: the operator corresponds to an observable and the outcome of a single act of detection is one of its eigenvalues. The only difficulty with the generalization is a technical one: one has to be sure that the eigenstates form a complete set.

•**symmetric operators with a continuum spectrum** There are situations where the problem $\hat{A}f=\lambda f$ cannot be solved *exactly* by square integrable functions, but only in the approximate sense of Weyl, i.e. there is a (Weyl) sequence $(f_n)_0^\infty$ of square integrable functions such that

$$\|\hat{A}f_n-\lambda f_n\|<\epsilon$$

for a sufficiently large n. Values of λ obtained in this way occupy typically some continuous interval and build the continuum spectrum. As an example of this situation we consider the symmetric operator $\hat{p}$. The eigenvalue equation for $\hat{p}$ reads

$$-i\frac{d}{dx}f(x) = \underbrace{\frac{p}{\hbar}}_{\doteq k}\cdot f(x)$$

It has the solutions

$$u_k(x) = \sqrt{\frac{1}{2\pi}}\cdot e^{i\cdot k\cdot x}$$

to any real value of k. These functions are not square integrable. Yet, one can think of using them within an integral to express any square integrable function according to

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\cdot\tilde{\psi}(k)\cdot e^{ikx} \quad \tilde{\psi}(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\cdot\psi(x)\cdot e^{-ikx}$$

The orthonormality relation writes

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\cdot k'\cdot x}\cdot e^{i\cdot k\cdot x} = \delta(k-k')$$

and the completeness relation writes (Plancherel)

$$\int_{-\infty}^{\infty} dx \mid \psi(x)\mid^2 = \int_{-\infty}^{\infty} dk \mid \tilde{\psi}(k)\mid^2$$

This example shows that on one side we have technically simple solutions at disposal, which are not really acceptable because not square integrable. On the other we have Weyl sequence approach, that tells us of the existence of square integrable functions that are acceptable but solve the problem only approximately and are not really useful. P.A.M. Dirac (and later rigorously L. Schwartz and I.M. Gel'fand) has provided us with the simple algorithm explained above, where a non square integrable function is incorporated into an integral containing a square integrable function to produce finite and useful results.

COMMENTS.

1. Regarding the outcome of an experiment aiming at measuring the momentum of the particle, we point out that the set of "eigenfunctions" is a complete one (by virtue of the theory of Fourier integrals). However, the possible eigenvalues occupy a continuous range and the question about the result of a single act of detection must be formulated accordingly. Given a certain wave function ψ, we would like to compute the probability to find a value of the momentum in a certain interval Δ. We compute this probability by generalizing the expression for the discrete spectrum:

$$\sum_{k:\lambda_k\in\Delta,}\sum_{n=1}^{n_k}\mid(f_{k,n},e)\mid^2$$

to

$$\int_{\Delta} dk\cdot\rho(k)\mid(u_k,\psi)\mid^2 = \int_{\Delta} dk\cdot\mid\tilde{\psi}(k)\mid^2$$

($\rho(k)$ is a "weight" that takes into account the degeneracy of the eigenvalues and their density in the interval dk. In the present case $\rho(k)$is just 1). This results is Born postulate for the variable "k" computed on the base of the spectral decomposition of the operator $\hat{k}$.

2. A further example of symmetric operator with continuous spectrum is the position operator $\hat{x}$. The eigenvalue equation

$$\hat{x}f(x) = \lambda f(x)$$

is solved, for any real λ, by

$$f(x) = \delta(x-\lambda)$$

This functions are not square integrable but can be used to express any square integrable function when incorporated within an integral, by virtue of the definition of the Dirac delta function. Accordingly, given a wave function ψ, the probability of finding a value of the position operator in a given interval Δ amounts to

$$\int_{\Delta} d\lambda \underbrace{|\int dx' \delta(x'-\lambda)\cdot\psi(x')\,|^2}_{|\psi(\lambda)|^2}$$

This result is Born's postulate for the variable "x" computed on the basis of the spectral decomposition of the operator $\hat{x}$.

•**Self-adjoint operators.** The anomalies observed in an infinite dimensional Hilbert space when dealing with symmetric operators show that being a symmetric operator is a necessary condition for representing an observable (because they have real eigenvalues), but it is not a sufficient one. The spectral theorem by Hilbert and v. Neumann finds exactly that subset of the symmetric operators that have a complete spectral decomposition in an infinite dimensional Hilbert space and correspond therefore to observables: the so called **self-adjoint operators**.

For the purpose of defining this class of operators we consider again, as an example, the differential operator

$$\hat{k} \doteq -i\frac{d}{dx}$$

and search first for those boundary conditions that render it symmetric.

$$\begin{aligned}\left(f,\hat{k}g\right) &= \int_{-\frac{L}{2}}^{\frac{L}{2}} dx\bar{f}(x)\cdot[-i\frac{d}{dx}g(x)] \overbrace{=}^{\text{P.}I.} -i\bar{f}(x)\cdot g(x)\,|_{-\frac{L}{2}}^{\frac{L}{2}} + \int_{-\frac{L}{2}}^{\frac{L}{2}} dx\overline{(-i\frac{d}{dx}f(x))}\cdot g(x)\\ &= -i\bar{f}(x)\cdot g(x)\,|_{-\frac{L}{2}}^{\frac{L}{2}} + \left(\hat{k}f,g\right)\end{aligned}$$

The operator can be rendered symmetric by e.g. letting it act only onto functions f,g for which $g(\pm\frac{L}{2})=0$ and $f(\pm\frac{L}{2})=0$. However, the boundary conditions for the function f, i.e. when the operator is on the left-hand side, need not to be the same as for the function g, as the partially integrated part vanishes even if $g(\pm\frac{L}{2})\neq 0$. The operator acting on the right hand side and on the left hand side are not precisely the same, as the equality

$$\left((\hat{k})_L f,g\right)=\left(f,(\hat{k})_R g\right)$$

can have solutions f,g which have different boundary conditions. The operator $(\hat{k})_L$ that fulfills the equation is called the operator adjoint to $\hat{k}$ and designated as $\hat{k}^{\dagger}$. It is formally the same as $\hat{k}$ but it has a different domain of definition. When the boundary conditions are chosen so that $\hat{k}^{\dagger}$ is defined exactly for the same functions as $\hat{k}$ and **no more**, then the operator $\hat{k}$ becomes a so called self-adjoint operator. $\hat{k}$ is e.g. self-adjoint if one choses periodic boundary conditions: under this assumption, the integrated part vanishes only if both set of functions f and g have periodic boundary conditions.

DEFINITION. The operator fulfilling the equation

$$\left(\hat{X}f,g\right)=\left(f,\hat{A}g\right)$$

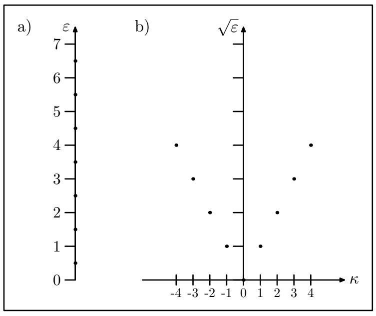
is called the adjoint of $\hat{A}$ and labeled $\hat{A}^{\dagger}$. If $\hat{A}^{\dagger}$ is formally identical to $\hat{A}$ and has the same domain of definition as $\hat{A}$ (and not larger) than $\hat{A}$ is a self adjoint operator.

THEOREM (Hilbert and v. Neumann): Self-adjoint operators have a complete spectral decomposition, i.e. have a complete set of eigenvalues and eigenfunctions and are, accordingly, the observables of quantum mechanics.

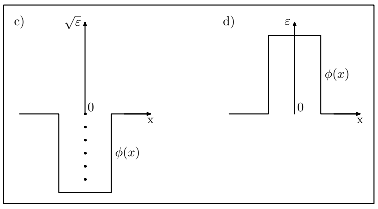
COMMENTS.

1. T. Kato, Trans. Amer. Math. Soc. **70**, p.195 (1951) has provided the proof that the Hamiltonian of an atom or a molecule is self-adjoint.

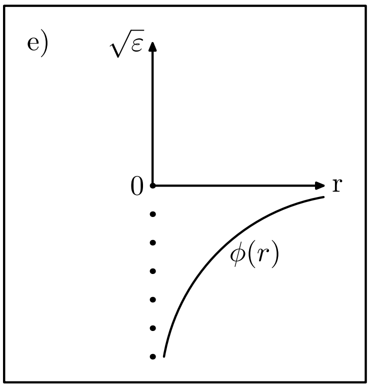
2. The most general situation of self-adjoint operators is that their eigenvalues assume discrete values within some range in $\mathscr{R}$ and continuous values within a different range. The following figures provides example of the various situations one can encounter, depending on the potential.



a: the spectrum of the quantum mechanical harmonic oscillator: the energy ϵ *is plotted in units of* $\hbar\omega$*. The spectrum (the dots along the vertical axis) is purely discrete. b: the spectrum of a particle moving in one dimension with periodic boundary conditions (period L). The square root of the energy* $\sqrt{\epsilon}$ *is plotted in units of* $\sqrt{\frac{\hbar^2}{2m}\cdot(\frac{2\pi}{L})^2}$ *as a function of* $\kappa = \frac{k}{\frac{2\pi}{L}}$.



c: The spectrum of a particle moving in a one dimensional potential well, typical of artificial atoms. The square root of the energy $\sqrt{\epsilon}$ *is plotted schematically. For negative energies the spectrum consists of discrete levels. For positive energies all energy values are possible and the spectrum builds a continuum. d: The spectrum of a particle moving in a one dimensional potential step. The spectrum has only a continuum component.*



e: The spectrum of a particle moving in a spherically symmetric potential (sketched). The spectrum has a discrete component for negative energies. The eigenvalue accumulate in the vicinity of $E = 0$*. For positive energies the spectrum has a continuum of values.*

3. For the general case of operators with discrete and continuous spectra, any wave function of the Hilbert space can be expanded as a sum of two components. One component consists of the **sum** over the eigenstates of the discrete spectrum. The second component is the **integral** over the (non-square integrable) basis states belonging to the continuum spectrum.
4. The outcome of an experiment will be either one of the eigenvalues in the discrete sector or some interval of values within the continuum spectrum. The probabilities are computed either as the sum or the integral of the absolute square of the suitable expansion coefficients. We will give below, as an optional part, the mathematical details of points 3 and 4.

Optional: the spectral decomposition for operators with discrete AND continuous eigenvalues. Suppose that we have a self-adjoint operator $\hat{A}$ and that the eigenvalue equation

$$\hat{A}\psi = \lambda\psi$$

has produced discrete eigenvalues λ_n and eigenfunctions f_n in some interval Δ_d with

$$(f_n, f_{n'}) = \delta_{nn'}$$

In addition, we assume a continuous interval of eigenvalues λ and (non-square integrable) eigenfunctions f_λ in the interval $\Delta_c = [a, b]$ (a, b can assume infinite values), with

$$(f_\lambda, f_{\lambda'}) = \int_{\mathcal{K}} d\Omega_q \overline{f_\lambda(q)} \cdot f'_\lambda(q) = \delta(\lambda - \lambda')$$

$\delta(\lambda - \lambda')$ is the Dirac-delta function. There is possibly an extra discrete index $m(n)$ that must be introduced to take into account the degeneracy of a discrete eigenvalue λ_n and a continuous (or discrete) density $\rho(\lambda, \nu)$ which is required for taking into account the degeneracy and the density of the eigenvalue λ. In the following, sum and integrals must be intended to include the index m and the density $\rho(\lambda, \nu)$.

- PROPOSITION: f_n and f_λ are orthogonal to each other.

PROOF. By definition, we have $\hat{A}f_n = \lambda_n f_n$ and $\hat{A}f_\lambda = \lambda_\lambda f_\lambda$. We choose a function

$$g(q) = \int_a^b d\lambda \tilde{g}(\lambda) \cdot f_\lambda(q)$$

i.e. we include the non-square integrable function into a "test integral": $\tilde{g}(\lambda)$ is any "test function" (a concept borrowed from the theory of distributions), i.e. a smooth function such that the "test integral" exists. We now compute. On one side

$$\begin{aligned} \left(f_n, \hat{A}g\right) &= \left(f_n, \hat{A}\int_a^b d\lambda \tilde{g}(\lambda) \cdot f_\lambda\right) \\ &= \int_a^b d\lambda \tilde{g}(\lambda) \cdot \lambda \cdot (f_n, f_\lambda) \end{aligned}$$

On the other

$$\left(f_n, \hat{A}g\right) = \lambda_n \cdot \int_a^b d\lambda \tilde{g}(\lambda) \cdot (f_n, f_\lambda)$$

Assuming that $\lambda_n \notin [a, b]$, the right hand side of both equations can only be equal, for any $\tilde{g}(\lambda)$, if $(f_n, f_\lambda) = 0$.

According to this result, the Hilbert space consists of the orthogonal sum of two subspaces: the one carrying the discrete spectrum and the one carrying the continuous spectrum. The expansion theorem reads therefore

$$f(q) = \sum_{n:\lambda_n \in \Delta_d} \sum_{m=1}^{d_n} \left(f_{n,m}, f\right) f_{n,m}(q) + \int_a^b d\lambda \int d\nu \rho(\lambda, \nu) \left(f_{\lambda,\nu}, f\right) \cdot f_{\lambda,\nu}(q)$$

The completeness relation writes

$$\| f \| = \sqrt{\sum_{n:\lambda_n \in \Delta_d} \sum_{m=1}^{d_n} |\left(f_{n,m}, f\right)|^2 + \int_a^b d\lambda \int d\nu \rho(\lambda,\nu) |\left(f_{\lambda,\nu}, f\right)|^2}$$

Finally the probability of finding one of the eigenvalues of $\hat{A}$ within the interval Δ_d of the discrete spectrum amounts to (we include the index m and the density ρ for completeness of results)

$$\sum_{n:\lambda_n \in \Delta_d} \sum_{m=1}^{d_n} |(f_{nm}, f)|^2$$

and the probability of finding one of the eigenvalues of $\hat{A}$ within the interval Δ_c of the continuum spectrum amounts to

$$\int_{\Delta_c} d\alpha \int d\nu \cdot \rho(\alpha,\nu) \cdot |\left(f_{\alpha,\nu}, f\right)|^2$$

The existence of a complete set of eigenvalues and eigenfunctions allows to compute functions of self-adjoint operators. From

$$\hat{A}\psi = \sum_{n,m} \lambda_n \cdot \left(f_{n,m}, \psi\right) f_{n,m} + \int_{\Delta_c} d\lambda \int d\nu \rho(\lambda,\nu) \cdot \lambda \cdot \left(f_{\lambda,\nu}, \psi\right) f_\lambda$$

one defines $F(\hat{A})$ as

$$F(\hat{A}) \doteq \sum_{n,m} F(\lambda_n) \cdot \left(f_{n,m}, \psi\right) f_{n,m} + \int_{\Delta_c} d\lambda \int d\nu \rho(\lambda,\nu) \cdot F(\lambda) \cdot \left(f_{\lambda,\nu}, \psi\right) f_{\lambda,\nu}$$

4.5. QM 2.0: Reduction of the state vector.

We now turn to a crucial question concerning the process of measurement of an observable A. The measurement process involves the coupling between a macroscopic device used to perform the measurement and a system which is described by the state ψ. After the measurement, the device has changed its state: we use this change to read out the value of the observable under detection. The question now is: given the fact that the macroscopic device has changed its state, what happens to the original state ψ itself?

This question is currently a much debated one. The determination of the eigenvalues of the Hamilton operator was at the very origin of the quantum mechanics. The finding of the energy levels in atoms, molecules and solids followed the next 50 years. Currently, very refined methods of computational physics help solving the eigenvalue problem of the Hamilton operator of complex materials and devices. We can call this aspects of quantum mechanics as QM 1.0. There is however a modern research topic – quantum technology (QM 2.0). Quantum technology is about testing the postulate that the result of one single act of measurement is one of the eigenvalues. It is also about finding an answer (both theoretical and experimental) to the question that inspires this section.

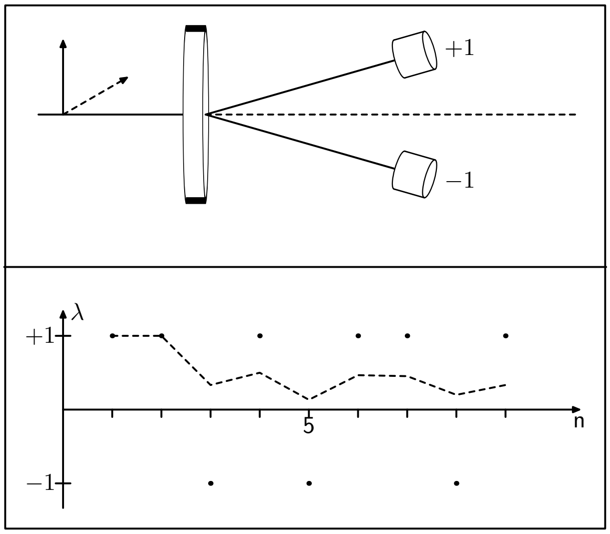
The origin of QM 2.0 is actually the answer provided by Dirac and von Neumann about the fate of the quantum mechanical state as a result of one single act of detection ; the "projection (or "reduction" or "jump") postulate. To properly formulate this postulate, we design an experiment involving a photon propagating along the z-direction that can be prepared in some state of polarization

$$|\psi> = a \cdot |x> + b \cdot |y>$$

The basis states $|x>$ and $|y>$ refer to states of the photon where its electric field is along the x-direction respectively y-direction. $|x>$ and $|y>$ can be considered as the eigenstates of a polarization operator $\hat{P}$ to the eigenvalues ±1.

$$\hat{P} \mid x >= +1 \mid x >$$
$$\hat{P} \mid y >= -1 \mid y >$$

The instrument to determining the observable P consists of a Wollastone prism placed along the path of the photon.



Top: A Wollastone prism directs photon with different polarisation along different paths. When they impact the counter, a signal is sent that writes a "+1" or "-1" into a register, depending on which counter has recorded the impact. Bottom: a possible sequence of measurements consists of "+1" or "-1". The dashed line gives the expectation value, which is a number in the interval $[-1, +1]$.

A Wollaston prism is a set of crystals that have the property of sending light polarized along x and y along a different path – say upwards for x and downwards for y. The two spatially separated beams are directed toward a counter that detects the arrival of light. When the upper counter register a photon, it writes the eigenvalue +1 on a register, when the lower counter register a photon it writes the eigenvalue −1. Let us now send a photon toward the measurement set up which we have prepared in the state $\mid \psi >$. The photon will be registered by one of the counters and the corresponding eigenvalue will be written down. We then repeat the experiment after having prepared the photon in the same state $\mid \psi >$. One of the counters will register a beep and will write down the corresponding eigenvalue. We repeat the measurement again and again – say a large number N and we might find a sequence of outcomes that are distributed between +1 and−1 exactly - no other outcome is expected. The question now is? How many times will we register +1, how many time will we register −1? The lore of quantum mechanics tells us the answer: the probability of measuring the eigenvalue +1 is namely $\mid a \mid^2$, the probability of measuring −1 is $\mid b \mid^2$. Of the N outcomes sampled by our instrument, $N_+ = \mid a \mid^2 \cdot N$ will be " plus one's", $N_- = \mid b \mid^2 \cdot N$ will be "minus one's". The quantum mechanical expectation value is some number lying in the interval $[-1, 1]$ and given by

$$< P > |_{\mid \psi >} = \frac{(+1) \cdot N_+ + (-1) \cdot N_-}{N_+ + N_-} = \frac{\mid a \mid^2 - \mid b \mid^2}{\mid a \mid^2 + \mid b \mid^2}$$

The experiment can also be performed with a stationary collimated beam of non-interacting photons with almost the same state of polarization $\mid \psi >$ such as the beam delivered by common laser sources: this would allow to measure the current of photons arriving at the two detectors almost instantaneously. The ratio of the two currents should be very much $\frac{|a|^2}{|b|^2}$.

There is a further aspect of this experiment that can be observed when a polarization analyzer it inserted between prism and detector. A polarization analyzer is a crystal with a preferred orientation that absorbs photon with a certain polarization. If we insert such an analyzer between the prism and the upper detector, oriented in such a way that photon with x polarization are completely absorbed, one registers a very peculiar outcome: the current arriving to the upper detector, that typically detect x-polarized photon, vanishes completely. This means that the Wollaston prism not only splits the beam into two components with different intensity, but is also changes the state $|\psi>$ of the incoming photon, projecting it into either one of the eigenstates. This is expression of a more general principle related to the measuring process in quantum mechanics (von Neumann:) "when an observable A is measured, the outcome of the measurement is one of the eigenvalues of $\hat{A}$. Simoulteneously, the state of the system is changed, and it is projected into the corresponding eigenspace". And Dirac wrotes: " From physical continuity, if we make a second measurement of the same observable $\hat{A}$ immediately after the first, the result of the second measurement must be the same as that of the first. Hence, after the first measurement has been made, the system is in an eigenstate of the dynamical variable $\hat{A}$, the eigenvalue belongs to being equal to the result of the first measurement. This conclusion must still hold if the second measurement is not actually made. In this way we see that the system JUMPS into an eigenstate of the observable that is being measured, the eigenvalue this eigenstate belongs to being equal to the result of the measurement. "
If the observable has a continuum spectrum, the result of the outcome of measuring the observable will be in some interval Δ, and the original state will be reduced to a superposition of the eigenstates belonging to Δ. For instance, when a positional detector has been placed to register the position of particle in some interval Δ, then the particle will consist of a wave function which is localized in the interval Δ.

EXAMPLES.
1. A particle with mass m in a segment of length L with periodic boundary conditions. A CONS consists of the plane waves

$$\frac{1}{\sqrt{L}} e^{i \cdot k \cdot x}$$

with $k = \frac{2\pi}{L} \cdot n \in \mathscr{Z}$ and

$$\int_0^L dx \frac{1}{\sqrt{L}} e^{-i \cdot k \cdot x} \frac{1}{\sqrt{L}} e^{i \cdot k' \cdot x} = \delta_{kk'}$$

The amplitude for the particle with wave-number k to be in the position x is given by

$$\psi_k(x) = \frac{1}{\sqrt{L}} \cdot e^{i(k \cdot x)}$$

Its probability density of being in the interval $[x, x+dx]$ is given by

$$|\psi_k(x)|^2 = \frac{1}{L}$$

regardless of x! If a segment δ is filled with a particle-tracer, one can ask about the probability of registering the particle within δ. Remember: the particle is a "point" (except for some inner structure that we do not consider here). It moves around in the plane wave state $\frac{1}{\sqrt{L}} e^{i \cdot k \cdot x}$. Then

$$W_{\hat{x}}^{\psi}(\delta) = \int_\delta dx \,|\, \frac{1}{\sqrt{L}} e^{i \cdot k \cdot x} \,|^2 = \frac{\delta}{L}$$

It follows that all positions x are equally likely: the state $\frac{1}{\sqrt{L}} e^{i \cdot k \cdot x}$ represents a so-called "completely

delocalized" state, where the particle can be detected everywhere with the same probability.
2. After performing the position measurement, the particle is localized within the segment δ_x. For describing this state, we use a simple function with

$$\psi(x)=\frac{1}{\sqrt{\delta_x}} \quad x\in[x-\frac{\delta_x}{2},x+\frac{\delta_x}{2}]$$

and

$$\psi(x)=0$$

elsewhere. We ask ourselves: what is the probability of measuring values of k in some interval δ_k? To answer this question we compute (let think of $L\rightarrow\infty$)

$$\psi(k)=\int_{x-\frac{\delta_x}{2}}^{x+\frac{\delta_x}{2}} dx'\frac{1}{\sqrt{2\pi}}\cdot e^{-i(k\cdot x')}\frac{1}{\sqrt{\delta_x}}=-\frac{1}{\sqrt{2\pi\cdot\delta_x}}\cdot 2\cdot e^{-i\cdot k\cdot x}\cdot\frac{\sin(k\cdot\frac{\delta_x x}{2})}{k}$$

The probability density amounts to

$$|\psi(k)|^2=\frac{4}{2\pi\cdot\delta_x}\cdot\frac{\sin^2(k\cdot\frac{\delta_x}{2})}{k^2}$$

The k-dependence is an oscillating function concentrated around $k=0$ and decaying as $\frac{1}{k^2}$.
If we define the width of the function in the k variable as, approximately, the distance between the two zero's closer to $k=0$, we find an interesting relation. As the zero's of this function are given by $\pi\cdot n\frac{2}{\delta_x}$, the width of the function amounts to $\delta_k=\frac{4\pi}{\delta_x}$. The width of the probability density in the variable x is given by δ_x. If we use the De Broglie hypothesis to rewrite k as $\frac{p}{\hbar}$ we can formulate the relation between δ_x and δ_p as follows:

$$\delta_x\cdot\delta_p=4\pi\cdot\hbar$$

The constant on the right hand-side is independent of the width of the wave functions. We therefore can change the width δ_x and we will observe that the width δ_p must change according to our relation! The smaller is δ_x (the width in the direct space) – i.e. the sharper the particle is localized in space, – the larger is δ_p – i.e. the broader is the wave function in p-space. In other words: trying to "localize" a particle in a narrow spatial area δ_x is only at the cost of greater uncertainty about its momentum, which can assume any value in the interval $[-\frac{2\pi\cdot\hbar}{\delta_x},\frac{2\pi\cdot\hbar}{\delta_x}]$.

Let us quote some sentences by W. Pauli (Pauli lectures on Physics, Dover Editions, Volume 5, ISBN 0-486-41462-0) on this issue. "If we consider a current of particles which passes through two small holes in a diaphragm, then we discover that the probability of finding a particle behind the diaphragm is a typical diffraction pattern, as in optics. This interference of the probability is independent of the intensity, that is independent of the density of particle current. It depends only on the locations of the holes; the holes define for us a state". One could continue: when a spatially localized state is defined, the momentum aquires a certain spread and the particles, afterwards, are redirected in space to follow a number of trajectories that are different from the original one and give rise to a "diffration pattern". "If we consider the state represented by a single, force-free atom, we find that this state is completely changed by every measurement of position or momentum. If we measure the position of the particle, for example, then an indeterminable momentum will be transferred to the particle during the process; this will make a precise determination of the position of the particle at a later time impossible. Although the trajectory of a celestial body can be determined more and more accurately by marking successive measurements to which the laws of classical mechanics are applied, every measurement on an elementary particle throws the particle out of its trajectory; that is, earlier measurements of position are useless for further determination of the trajectory".

The equation

$$\delta_x \cdot \delta_p = 4\pi \cdot \hbar$$

has been obtained assuming a well-definded shape of the wave function. It is a special case of the famous **uncertainty relation of Heisenberg**:

$$(\delta_x)\cdot\left(\delta_p\right) \geq \frac{\hbar}{2}$$

The $=$ sign is realized when the wave function in real space has a Gaussian profile.

COMMENT.
1. The Heisenberg uncertainty relation can be generalized to any pair of observables, as shown by the following
THEOREM. Let $\hat{A},\hat{B}$ be to two operators and define for any operator the square of the standard deviation

$$(\Delta_\psi \hat{A})^2 \doteq <\left(\hat{A} - <\hat{A}>_\psi \mathbb{1}\right)^2 >_\psi = <\hat{A}^2>_\psi - <\hat{A}>_\psi^2$$

The following inequality holds true (Robertson inequality)

$$(\Delta_\psi \hat{A})^2 \cdot (\Delta_\psi \hat{B})^2 \geq \frac{1}{4} |<[\hat{A},\hat{B}]>_\psi|^2$$

where the operation $[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ is the so called ”commutator” of the operators $\hat{A},\hat{B}$. As an example:

$$[\hat{x},\hat{p}] = i\hbar \cdot \mathbb{1}$$

Inserting this result into the Robertson inequality gives the Heisenberg uncertainty principle.

- **Commuting operators.** The Robertson inequality for commuting operators simplifies to,

$$(\Delta_\psi \hat{A})^2 \cdot (\Delta_\psi \hat{B})^2 \geq 0$$

This means that when a certain value of the observable A is detected and the original state is reduced to an eigenspace of $\hat{A}$ ($\Delta\hat{A} = 0$), a successive measurement of the observable B can also give a sharp value, i.e. $\Delta\hat{B}$ can be also vanishing.

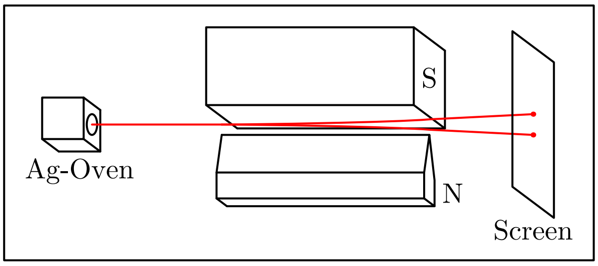
5. The Spin, the algebra of the angular momentum and the merging of degrees of freedoms.

5.1. 1925: the Stern-Gerlach experiment and the spin hypothesis.

The states originating within the SE for a spherical potential have following properties:

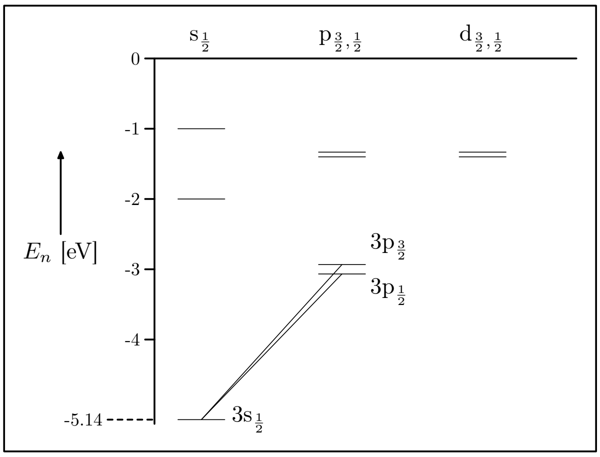
- $\mathbf{L}^2 Y_l^m = \hbar^2 \cdot l \cdot (l+1) \cdot Y_l^m$, $l = 0,1,2,\ldots$
- $L_z Y_l^m = \hbar m \cdot Y_l^m$, $m = l, l-1, \ldots, -l$
- $\{Y_l^m\}$ (Dirac: $\{|\, l, m >\}$) build a $2l+1$-dimensional space, i.e. the degeneracy of the eigenvalue $\hbar^2 \cdot l \cdot (l+1)$ is $2l+1$. Moreover $\left(Y_l^m, Y_{l'}^{m'}\right) = \delta_{ll'} \cdot \delta_{mm'}$
- l is the orbital angular momentum quantum number, m is the magnetic quantum number
- Y_l^m is a state that can be specified by giving its (possibly complex valued) wave function defined onto the unit sphere at the spherical coordinates ϑ, φ i.e. $Y_l^m(\vartheta, \varphi)$ (we have seen some of these wave functions). Dirac:
$$|\, l, m > \quad \rightarrow < \vartheta, \varphi \,|\, l, m >$$
- The above equalities can be obtained by applying the differential operators $\mathbf{L}^2(\vartheta, \varphi)$ and $L_z(\varphi)$ onto the spherical harmonics $Y_l^m(\vartheta, \varphi)$.

The quantum numbers n, l are not enough for defining reliable basis states for the purposes of computing properties of atoms, molecules and matter. In 1921, O. Stern and W. Gerlach did a very remarkable experiment that added a new quantum number.



They placed at the entrance of an instrument that nowadays is known as the Stern-Gerlach apparatus an Ag furnace. If the furnace is heated, Ag atoms evaporate. At that time the Ag-atom was known to have one single electron circulating around a set of shells – to be indented as Bohr circular orbits – containing the remaining electrons. According to the quantization rule by Bohr, such an electron has angular momentum quantum number 1. In such a state, the z-component of the quantized angular momentum was expected to be either $\hbar$, 0 or $-\hbar$. In classical physics, all values in-between would also be possible. This means that the Ag-atoms were expected to leave the furnace equally distributed in some state of the z-component of the angular momentum – one of the three quantized states in the Bohr-Sommerfeld quantized model – or a continuous value in the interval $[-\hbar, \hbar]$ in the classical model. Such an experiment was therefore designed to discriminate between the two models. After leaving the furnace, the beam is introduced in an area where there is a strong magnetic gradient: $\mathbf{B}$ - predominantly in the z direction - is highly inhomogeneous along the z axis. Such inhomogeneity can be achieved by properly designing the poles of the magnet. It is known form classical electromagnetism that the magnetic field couples with the angular momentum (better expressed: with the magnetic moment associated with the angular momentum). The inhomogeneous magnetic field along z is able to send atoms with different z-component of the angular momentum along different path along the vertical z-axis. In other words: in the absence of magnetic field the Ag-atoms passes the region of B-inhomogeneity unperturbed and arrive at the center of the screen. When the magnetic field is on, two scenarios were expected: in the Bohr-quantized scenario three separate beams would form and produce three separate spots along z on the screen. In the scenario based on classical physics the particles would distribute themselves along a continuous line in the z-direction. Stern and Gerlach observed none of these two expected scenarios: they observed two separate spots. This was taken as partial proof of quantization of angular momentum, but led to speculations by e.g. Sommerfeld about the absence of the third spot. Could it be possible that, by some reason, one of the three beams was somehow ”suppressed”? The correct interpretation of this result was formulated as an hypothesis by Goudsmit and Uhlenbeck in 1925: the electron is actually in a state of vanishing orbital angular momentum (s-state, as we know from the Schrödinger quantum mechanics) but must have had some internal degree of freedom with angular momentum character and quantum number $\frac{1}{2}$ – in order to explain the existence of only two z-components. This internal degree of freedom was called the ”spin” of the electron.

The hypothesis by Goudsmit and Uhlenbeck was based on further observations that pointed to a new double valued quantum number: in the same year 1925 Pauli suggested his principle of double occupancy of the Bohr circular orbits – suggesting the existence of a new double valued quantum number. In the meantime there were spectroscopic observations on the Na vapor. The transition drawn in the figure is the well-known Na-D line ($\lambda \approx 589$ nm), which causes a yellow colour of the Na-flame. At sufficiently high resolution of the spectrometer it can be seen that it consists of two closely spaced lines - a so-called doublet (fine structure of the spectral lines, a term introduced by A. Sommerfeld). Again, this observation pointed to a new, doubled-valued quantum number for the angular momentum.



Term-scheme of the Na-atom indicating the transitions producing the "Na-doublet". The energy levels are labeled with the main quantum number 3, the orbital quantum number 1 and a third quantum number referring to the quantum number of the total angular momentum – orbital plus spin quantum number. This third quantum number will be explained later.

COMMENTS.

1. The existence of half-integer spin quantum numbers raises the question of the existence of $S = \frac{1}{3}, S = \frac{1}{4}$, and so on. The charge, e.g., occurs in integer multiples of $|\, e \,|$, but $\frac{1}{3} \,|\, e \,|$ exists too! The answer to this question is provided by general symmetry arguments: it is the rotational symmetry of the space that allows only integer and half-integer angular momenta.

2. The Schrödinger equation foresees that quantum numbers for orbital angular momenta are integer numbers $l = 0,1,2,3,...$. Half-integer numbers can only be assigned to the intrinsic spin state of particles. Of course, there are particles that have an integer intrinsic spin state quantum numbers, such as photons which have spin 1 and gravitons (spin 2).

3. Orbital angular momentum quantum numbers are often given with the Latin letter small l (or large L). Intrinsic spin angular momentum quantum numbers are often given with the small Latin letter s or large S. Quantum numbers resulting from both (see later for how to add angular momenta) are given with the small Latin letter j or large J.

4. The angular momentum operators defined as differential operators in the variables ϑ, φ produce integer orbital quantum numbers l and integer magnetic quantum numbers $m = l, l-1, ..., -l$. The eigenstates are described by the spherical harmonics $\{Y_l^m(\vartheta, \varphi)\}$. The spherical harmonics give the amplitude that a particle with angular quantum numbers l, m is found along the spatial direction specified by the angles (ϑ, φ). The configuration space for specifying the states of the orbital angular momentum is the unit sphere. For the operators of orbital angular momentum a representation in terms of differential operators over the variables ϑ and φ exists. For the operators with half integer spins there is no representation in terms of differential operators over the variables (ϑ, φ) and the eigenstates for half-integer spins cannot be specified by amplitudes that refer to angular variables. This is because when, e.g., a spin-state of an half-integer spin is rotated by 2π it gets multiplied by -1. Only rotations by 4π restore the original state. This specific rotational property cannot be described using the variables (ϑ, φ) on a unit sphere in Euclidean space: Non-integer spin states do

not have a configuration space residing in Euclidean space.
The question is now: how do we define operators and states for the non-integer spin angular momenta? The answer is

DEFINITION. One defines basis states Y_s^m, $m = s, s-1, ... -s$ for half integer spins s (so called spinors) and spin operators $\hat{S}_x$, $\hat{S}_y$, $\hat{S}_z$ by assuming that these abstract items obey the same computations rules obtained in the Schrödinger quantum mechanics for integer l.

For instance,

$$\hat{S}_z Y_s^m = \hbar m \cdot Y_s^m \quad \hat{\mathbf{S}}^2 Y_s^m = \hbar^2 \cdot s(s+1) Y_s^m$$

For the particular case of $s = \frac{1}{2}$ one defines the basis states $Y^{\pm}$:

$$\hat{S}_z Y^{\pm} = \pm\hbar \frac{1}{2} \cdot Y^{\pm} \quad \hat{\mathbf{S}}^2 Y^{\pm} = \hbar^2 \cdot \frac{3}{4} Y^{\pm}$$

Further symbols for these basis states are $\psi_{\pm}$, $\psi_{\uparrow\downarrow}$, $|\uparrow>$ and $|\downarrow>$ oder $|\,\mathrm{up}>$ and $|\,\mathrm{down}>$ and so on. The Hilbert space of for the internal spin-s degree of freedom is labeled as Θ_s.

5. One can obtain the description of a spinor as a $\mathscr{C}$-valued scalar field by introducing a configuration space $\mathscr{K}_s = \{+s, ..., -s\}$ for the variable m and a map

$$Y_s : m \in \mathscr{K}_s \to Y(m) \in \mathscr{C}$$

$Y_s(m)$ is the amplitude for the spin to assume the value m of the z-component of the spin angular momentum. In this notation the z-component of the spin angular momentum works as an independent variable over which the wave function is described. For the basis states $Y_s^{m_s}$, $m_s = +s, ..., -s$ we write

$$Y_s^{m_s}(m) = \delta_{m,m_s}$$

We have the computation rules:

$$\hat{S}_z Y_s^{m_s}(m) = \hbar m_s \cdot Y_s^{m_s}(m) \quad \hat{\mathbf{S}}^2 Y_s^m(m) = \hbar^2 \cdot s(s+1) Y_s^{m_s}(m)$$

In summary: a spinor can be written as a sum of basis states

$$Y_S = \sum_{m=-S}^{S} c_m \cdot Y_S^m$$

or as a sum of basis functions

$$Y_S(m) = \sum_{m_S} c_{m_S} \cdot Y_S^{m_S}(m)$$

5.2. The algebra of angular momentum: a summary

The orbital angular momentum plays a crucial role in quantum mechanics because it appears as ”good” quantum number in many problems of atomic and particle physics. i.e it can be used to label the eigenstates. In addition, particles have not only an orbital angular momentum (as the one we have encountered in the term scheme of the Hydrogen atom) but have an intrinsic spin that, via Pauli principle, profoundly determine their physical behaviour: macroscopic states of matter like metals, semiconductors, magnetism, superconductivity, Bose condensation, all depend of the spin of the individual particles involved. The quantum mechanical orbital and spin angular momentum operator have a particular algebra that we need to learn about.
The path toward learning how to do computations with angular momenta starts with the expression that we have obtained within Schödinger representation. The representation of the operator of the

orbital momentum in spherical coordinates reads

$$\begin{aligned}\hat{L}_x &= i\hbar\cdot\left(\sin\varphi\frac{\partial}{\partial\vartheta}+\cot\vartheta\cos\varphi\frac{\partial}{\partial\varphi}\right)\\ \hat{L}_y &= -i\hbar\cdot\left(\cos\varphi\frac{\partial}{\partial\vartheta}-\cot\vartheta\sin\varphi\frac{\partial}{\partial\varphi}\right)\\ \hat{L}_z &= -i\hbar\frac{\partial}{\partial\varphi}\\ \hat{\mathbf{L}}^2 &= -\hbar^2\left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}(\sin\vartheta\frac{\partial}{\partial\vartheta}+\frac{1}{\sin\vartheta}\frac{\partial^2}{\partial\varphi^2}\right]\end{aligned}$$

When these operators are applied onto the spherical harmonics $Y_l^m(\vartheta,\varphi)$ – applied means that one operates onto the spherical harmonics by taking the appropriate derivatives with respect to appropriate variables, a set of results are produced that can be summarized as computation rules. As an example we apply $\hat{L}_z=-i\hbar\frac{d}{d\varphi}$ onto the generical spherical harmonic $Y_l^m(\vartheta,\varphi)$:

$$-i\hbar\frac{d}{d\varphi}\left[\underbrace{a_{lm}}_{\text{normalization constant}}\cdot\underbrace{P_l^m(\cos\vartheta)}_{\text{depends only on}\vartheta}\cdot e^{i\cdot m\cdot\varphi}\right]=$$

$$-i\hbar\cdot a_{l,m}\cdot P_l^m(\cos\vartheta)\cdot\frac{de^{i\cdot m\cdot\varphi}}{d\varphi}=\hbar\cdot m\cdot Y_l^m((\vartheta,\varphi)$$

What we have just computed is the computation rule

$$\hat{L}_z(\varphi)Y_l^m(\vartheta,\varphi)=\hbar\cdot m\cdot Y_l^m(\vartheta,\varphi)$$

In words: "the operator $\hat{L}_z$ lowers the upper index m of any spherical harmonics and places it as a multiplicative factor, together with $\hbar$ in front of the spherical harmonic itself."

Using the differential expressions for L_x, L_y and $\hat{\mathbf{L}}^2$ we can produce a set of such computation rules that concern all components of the orbital angular momentum operator and its square. For instance, by explicitly applying the differential operator for $\hat{\mathbf{L}}^2$ we obtain

$$\hat{\mathbf{L}}^2 Y_l^m(\vartheta,\varphi)=\hbar^2\cdot l(l+1)Y_l^m(\vartheta,\varphi)$$

In words: the operator of the square of the vector of the angular momentum takes the lower index of any spherical harmonics (independent on m) and puts it as a prefactor $l\cdot(l+1)$, together with $\hbar^2$ in front of the spherical harmonics itself. There is a main difference with $\hat{L}_z$: the index is l but the prefactor is not just $\hbar^2 l$but $\hbar^2 l\cdot(l+1)$. This is because the eigenvalues of $\hat{L}_z$ are related to the m-values (the magnetic quantum numbers), the eigenvalues of $\hat{\mathbf{L}}^2$ are not just the orbital angular momentum quantum numbers $\hbar^2 l$ but $\hbar^2 l\cdot(l+1)$. We can of course rewrite these computations rules in terms of the states Y_l^m:

$$\hat{L}_z Y_l^m=\hbar m\cdot Y_l^m \quad \hat{\mathbf{L}}^2 Y_l^m=\hbar^2\cdot l(l+1)Y_l^m$$

and produce similar computations rules for $\hat{L}_x$ and $\hat{L}_y$ (these are less frequently used). These computation rules must be supplemented by an important information: The eigenspaces to a given l are $2l+1$-times degenerate, corresponding to the $2l+1$ values for the eigenvalues of $\hat{L}_z$.

The values for l allowed for describing the orbital degree of freedom are, according to the Schrödinger quantum mechanics, strictly integers $0,1,2,\ldots.$ We have just found that half-integer angular momenta are also possible. Their computations rules can be obtained by translating those for integer l to half-integer values, although a differential operator for half integer spins does not exist. In the following, we give the full set of computation rules, valid for any angular momentum quantum number J – orbital or spin, integer or half-integer (after Condon, E. U. and Shortley, G. The Theory of Atomic Spectra. Cambridge, England: Cambridge University Press, 1951).

- $\hat{J}_x Y_j^m=\frac{\hbar}{2}\cdot\sqrt{j(l+1)-m(m+1)}\cdot Y_j^{m+1}+\frac{\hbar}{2}\cdot\sqrt{j(j+1)-m(m-1)}\cdot Y_j^{m-1}$

- $\hat{J}_y Y_j^m = -i\frac{\hbar}{2} \cdot \sqrt{j(j+1)-m(m+1)} \cdot Y_j^{m+1} + i\frac{\hbar}{2} \cdot \sqrt{j(j+1)-m(m-1)} \cdot Y_j^{m-1}$
- $\hat{J}_z Y_j^m = \hbar m \cdot Y_j^m$
- $\hat{\mathbf{J}}^2 Y_J^m = \hbar^2 \cdot J(J+1) Y_J^m$

COMMENT.
For spin $\frac{1}{2}$ particles there is a special terminology: one introduces a vector operator σ by the equation

$$\hat{\mathbf{S}} \doteq \frac{\hbar}{2}\sigma$$

σ is the so called Pauli operator.It is used for short writing $\hat{\mathbf{S}}$. We have

$$\begin{aligned} \hat{S}_x Y^{\pm} &= \frac{\hbar}{2} \cdot Y^{\mp} &\Leftrightarrow\quad& \sigma_x Y^{\pm} = Y^{\mp} \\ \hat{S}_y Y^{\pm} &= \pm i \cdot \frac{\hbar}{2} \cdot Y^{\mp} &\Leftrightarrow\quad& \sigma_y Y^{\pm} = \pm i \cdot Y^{\mp} \\ \hat{S}_z Y^{\pm} &= \pm\frac{\hbar}{2} \cdot Y^{\mp} &\Leftrightarrow\quad& \sigma_z Y^{\pm} = \pm Y^{\pm} \\ \vec{\hat{S}}^2 Y^{\pm} &= \hbar^2 \frac{3}{4} \cdot Y^{\mp} &\Leftrightarrow\quad& \sigma^2 Y^{\pm} = 3 \cdot Y^{\pm} \end{aligned}$$

EXAMPLES.
1. We use the computation rules to obtain the matrix representation of L_z and $(\mathbf{L})^2$ within the subspace provided by p-states. From

$$\mathbf{L}^2 Y_1^m = \hbar^2 \cdot 1 \cdot (1+1) \cdot Y_1^m$$

and

$$L_z Y_1^m = \hbar m \cdot Y_1^m$$

$(m = 1, 0, -1)$ we compute the matrix representations of L_z and $\mathbf{L}^2$ in the space (Y_1^1, Y_1^0, Y_1^{-1}). We need the nine matrix elements $\left(Y_1^m, L_z Y_1^{m'}\right)$ and the nine matrix elements $\left(Y_1^m, \mathbf{L}^2 Y_1^{m'}\right)$. We know that the operator L_z just lowers the upper index m', to produce the result $L_z Y_1^{m'} = \hbar \cdot m' \cdot Y_1^{m'}$. We then use the ortho-normality relations for spherical harmonics to obtain

$$\left(Y_1^m, L_z Y_1^{m'}\right) = \hbar \cdot m' \cdot \delta_{mm'}$$

This shows that the matrix representation of L_z in the subspace $\{Y_1^1, Y_1^0, Y_1^{-1}\}$ has only diagonal matrix elements:

$$L_z = \hbar \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{1} \end{pmatrix}$$

We compute now $\left(Y_1^m, \mathbf{L}^2 Y_1^{m'}\right)$. We know that $\mathbf{L}^2$ applied to $Y_1^{m'}$ just multiply $Y_1^{m'}$ with $\hbar^2(1(1+1))$, for any m'! This, together with the orthogonality relations, produce

$$\left(Y_1^m, \mathbf{L}^2 Y_1^{m'}\right) = \hbar^2 \cdot 1(1+1) \cdot \delta_{mm'}$$

This shows that the matrix representation of $\mathbf{L}^2$ in the subspace $\{Y_1^1, Y_1^0, Y_1^{-1}\}$ has only the identical diagonal matrix elements:

$$\mathbf{L}^2 = \hbar^2 \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

2. Here are some useful matrices for the operators $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ and ($\mathbf{L}^2$ in the subspaces of s, p and d electrons.

$$\underbrace{\hbar\cdot\begin{pmatrix}0\end{pmatrix} \quad ; \quad \hbar\cdot\begin{pmatrix}1&0&0\\0&0&0\\0&0&\bar{1}\end{pmatrix} \quad ; \quad \hbar\cdot\begin{pmatrix}2&0&0&0&0\\0&1&0&0&0\\0&0&0&0&0\\0&0&0&\bar{1}&0\\0&0&0&0&\bar{2}\end{pmatrix}}_{\hat{L}_z}$$

$$\underbrace{\hbar\cdot\begin{pmatrix}0\end{pmatrix} \quad ; \quad \hbar\cdot\frac{\sqrt{2}}{2}\cdot\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix} \quad ; \quad \hbar\cdot\begin{pmatrix}0&1&0&0&0\\1&0&\sqrt{\frac{3}{2}}&0&0\\0&\sqrt{\frac{3}{2}}&0&\sqrt{\frac{3}{2}}&0\\0&0&\sqrt{\frac{3}{2}}&0&1\\0&0&0&1&0\end{pmatrix}}_{\hat{L}_x}$$

$$\underbrace{\hbar\cdot\begin{pmatrix}0\end{pmatrix} \quad ; \quad \hbar\cdot\frac{\sqrt{2}}{2}\cdot\begin{pmatrix}0&-i&0\\i&0&-i\\0&i&0\end{pmatrix} \quad ; \quad \hbar\cdot\begin{pmatrix}0&-i&0&0&0\\i&0&-i\sqrt{\frac{3}{2}}&0&0\\0&i\sqrt{\frac{3}{2}}&0&-i\sqrt{\frac{3}{2}}&0\\0&0&i\sqrt{\frac{3}{2}}&0&-i\\0&0&0&i&0\end{pmatrix}}_{\hat{L}_y}$$

$$\underbrace{\hbar^2\cdot\begin{pmatrix}0\end{pmatrix} \quad ; \quad \hbar^2\begin{pmatrix}2&0&0\\0&2&0\\0&0&2\end{pmatrix} \quad ; \quad \mathbf{L}^2=\hbar^2\cdot\begin{pmatrix}6&0&0&0&0\\0&6&0&0&0\\0&0&6&0&0\\0&0&0&6&0\\0&0&0&0&6\end{pmatrix}}_{\hat{\mathbf{L}}^2}$$

3. Spin $\frac{1}{2}$: Referring to the basis states $\{Y^+, Y^-\}$, the matrix representation of the basis spinors is

$$\begin{pmatrix}1\\0\end{pmatrix}; \quad \begin{pmatrix}1\\0\end{pmatrix}$$

This representation establishes the isomorphism between $\Theta_{\frac{1}{2}}$ and $\mathscr{C}^2$. In general, the Hilbert space of a S-spinor Θ_S is isomorphic to $\mathscr{C}^{2S+1}$. The spin $\frac{1}{2}$ matrices reads

$$\hat{S}_z=\hbar\begin{pmatrix}\frac{1}{2}&0\\0&-\frac{1}{2}\end{pmatrix} \quad \hat{S}_x=\hbar\begin{pmatrix}0&\frac{1}{2}\\\frac{1}{2}&0\end{pmatrix} \quad \hat{S}_y=\hbar\begin{pmatrix}0&-\frac{i}{2}\\\frac{i}{2}&0\end{pmatrix}$$

and

$$\sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix} \quad \sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix} \quad \sigma_y=\begin{pmatrix}0&-i\\i&0\end{pmatrix}$$

5.3. Coupling of degrees of freedom in quantum mechanics

5.3.1. Introduction

For a complete description of a quantum mechanical particle the specification of its orbital angular momentum quantum number L is not sufficient: its spin S, if any exists, must be ”added” to it. This operation of ”adding angular momenta” is an example of a more general task that we have not yet discussed: the coupling of degrees of freedom in quantum mechanics.
Suppose a particle is moving along the x-axis in a state ψ, specified by the wave function $\psi(x \in \mathscr{R})$. Along y let the state φ be specified by $\varphi(y \in \mathscr{R})$. Suppose now to open both degrees of freedom to the particle. By this operation, the configuration space is enlarged from the x- and y -lines to the (x,y)-plane. Technically: the configuration space is the cartesian product $\mathscr{R} \times \mathscr{R}$, i.e. the particle coordinates become the pair (x,y). In classical physics, the joint system is described by enlarging the phase space to include both coordinates and both momenta. The question in quantum physics is about how the states of the joint system are described, given the knowledge of the states defined onto the two separate coordinates. Von Neumann and Dirac introduce, for this purpose, the operation of a ”product” of the states $|\psi>$ and $|\varphi>$. Exactly this product we would like to specify accurately. Notice that a product of states defined onto the same variable is not allowed in quantum mechanics: it violates the requirement of linearity. On the other side, we have encountered a product of states while solving the Schrödinger equation for the Hydrogen atom, in the form of a very important technical tool: the separation Ansatz. With this Ansatz we have created a trial wave function which consists of the product between wave functions specified on different spatial variables – in that specific case r, ϑ, φ. The separation Ansatz is the practical proof that a product of wave functions does not violate the superposition principle – provided the functions are defined onto different variables. Looking back at the physical significance of the separation Ansatz, a trial wave function consisting of a product gives the amplitudes to be at a certain radius r and along a certain direction specified by ϑ,φ. The product of amplitudes thus provides us with the correct probability amplitude for the realization of the joint variables r,ϑ,φ, given the amplitudes for the realization of the single variables. It makes therefore sense to introduce a product between states that is designed to realize the coupling of different degrees of freedom along the lines indicated by the separation Ansatz.

5.3.2. The Kronecker (tensor) product of states.

The general situation is one where the configuration space foresees two degrees of freedom q_1 and q_2. The variables q_i might be e.g. different spatial coordinates of one given particle – in this case each q_i takes a continuum set of values extending to infinity or, possibly, limited to some segment of finite length. q_i, however, can also give the coordinates of a further particle that is added to build a composite system. Or q_i can indicate the z-component of the spin, in which case q_i can take the $2S+1$ discrete values between $+S$ and $-S$. Let ψ be a state of a system with configuration space $\{q_1\}$ and φ the state of a system with configuration space $\{q_2\}$.

DEFINITION. The state $\psi \otimes \varphi$ is defined in such a way that, when evaluated at (q_1,q_2) it gives $\psi(q_1)\cdot\varphi(q_2)$, i.e.

$$(\psi \otimes \varphi)(q_1,q_2) = \psi(q_1)\cdot\varphi(q_2)$$

and is called the Kronecker (or tensor) product of the two states ψ and φ. The product on the right-hand side is the standard product between complex numbers.
The tensor product is a way of signalling that the two states being coupled contain variables that refer to **different** degrees of freedom. Dirac, starting from $|\psi>$ and $|\varphi>$, writes simply $|\psi>|\varphi>$

COMPUTATION RULES.

This definition that translates the tensor product of states onto the product of complex numbers implies following rules of computation:

a) $\otimes$ is linear in each slot in the sense that for any complex number α, β we have

$$(\alpha\psi_1+\beta\psi_2)\otimes\varphi=\alpha\psi_1\otimes\varphi+\beta\psi_2\otimes\varphi$$

and

$$\psi\otimes(\alpha\varphi_1+\beta\varphi_2)=\psi\otimes(\alpha\varphi_1)+\psi\otimes(\beta\varphi_2)$$

and

$$\alpha(\psi\otimes\varphi)=(\alpha\psi)\otimes\varphi=\psi\otimes(\alpha\varphi)$$

By these rules, the set of all linear combinations of $\{\psi\otimes\varphi\}$ becomes a vector space.

b) In this vector space a scalar product is defined by

$$(\psi_1\otimes\varphi_1,\psi_2\otimes\varphi_2)=(\psi_1,\psi_2)\cdot(\varphi_1,\varphi_2)$$

c) If $\{e_1,.....\}$ and $\{f_1,....\}$ are CONS in the respective Hilbert spaces, then the set $\{e_1\otimes f_1,e_1\otimes f_2,.....,e_2\otimes f_1,e_2\otimes f_2,......\}$ builds a CONS in the tensor product space, which also becomes an Hilbert space. Given the existence of a CONS, every state in the tensor product space can be expanded as linear combination of basis states:

$$\Psi=\sum_{i,j}c_{ij}\cdot e_i\otimes f_j$$

d) The scalar product rule $(\psi_1\otimes\varphi_1,\psi_2\otimes\varphi_2)$ and the CONS expansion lead to the general rule for the scalar product between states:

$$(\Psi,\Phi)=\int dq_1\int dq_2\,\overline{\Psi(q_1,q_2)}\cdot\Phi(q_1,q_2)$$

COMMENT.

Notice that the CONS expansion can be such that Ψ cannot be written as a product $\psi\otimes\varphi$: in this case, Ψ represents a so called **entangled** state. In other words, by means of this expansion one can create wave functions $\Psi(q_1,q_2)$ which cannot be written as the product $\psi(q_1)\cdot\varphi(q_2)$. Entangled states are the subject of modern research on quantum technology and quantum computing.

EXAMPLES.

1. The set

$$e^{i\frac{2\pi}{L}\cdot n\cdot x}$$

with $n\in\mathscr{Z}$ builds a CONS for the Hilbert space of periodic function in the variable x. Similarly, the set

$$e^{i\frac{2\pi}{L}\cdot m\cdot y}$$

with $m\in\mathscr{Z}$ builds a CONS for the Hilbert space of periodic function in the variable x. The Hilbert space of periodic functions over the configuration space $[0,L]\times[0,L]$ is spanned by the tensor products

$$\{e^{i\frac{2\pi}{L}\cdot n\cdot x}\cdot e^{i\frac{2\pi}{L}\cdot m\cdot y}\}=\{e^{i\frac{2\pi}{L}\cdot(n\cdot x+m\cdot y)}\}=\{e^{i\mathbf{kr}}\}$$

with $\mathbf{k}=\frac{2\pi}{L}(n,m)$ and $\mathbf{r}=(x,y)$.

2. The tensor product is used to construct the full Hilbert space of an electron with spin. The spatial degrees of freedoms are spanned by the Hilbert space $\mathscr{L}(\mathscr{R}^3)$. The Hilbert space describing the spin degree of freedom is Θ_2. The full Hilbert space of a spinning electron is

$$\mathscr{L}(\mathscr{R}^3)\otimes\Theta_2$$

Let $\{\psi_i\}$ build a CONS in $\mathscr{L}(\mathscr{R}^3)$. Then a spinor wave function writes

$$\begin{aligned}\psi(\mathbf{r},m) &= \sum_{i,+} c_{i,+}\psi_i(\mathbf{r})\cdot Y^+(m)+\sum_{i,-} c_{i,-}\psi_i(\mathbf{r})\cdot Y^-(m)\\ &= \psi_+(\mathbf{r})\cdot Y^+(m)+\psi_-(\mathbf{r})\cdot Y^-(m)\end{aligned}$$

or

$$\psi=\psi_+\otimes Y^+ +\psi_-\otimes Y^-$$

3. Consider a composite system consisting of two particles with orbital angular momentum quantum number L_1 and respectively $L_2 (L_i = 0,1,2,...)$. The $(2L_1+1)\cdot(2L_2+1)$ product functions

$$\underbrace{Y_{L_1}^{m_1}(\vartheta_1,\varphi_1)}_{2L_1+1}\cdot\underbrace{Y_{L_2}^{m_2}(\vartheta_2,\varphi_2)}_{2L_2+1}$$

with $m_L = L_1,...,-L_1$, $m_2 = L_2,...,-L_2$ build a subspace of $\mathscr{L}(S^2,d\Omega_1)\otimes\mathscr{L}(S^2,d\Omega_2)$. The set of all basis functions to any L_1 and L_2 build a CONS in $\mathscr{L}(S^2,d\Omega_1)\otimes\mathscr{L}_2(S^2,d\Omega_2)$ and can be used to expand periodic functions defined onto the four variables $(\vartheta_1,\varphi_1,\vartheta_2,\varphi_2)$.

5.3.3. The tensor (Kronecker) product of operators

We need to find out how operators act onto a tensor product state. For this purpose one defines the tensor product of operators by following operation:

DEFINITION. Let $\hat{A}_1$ and $\hat{A}_2$ be operators acting onto the Hilbert space ”1” and ”2” respectively. Then the tensor product

$$\hat{A}_1\otimes\hat{A}_2$$

is defined on a product of states by

$$(\hat{A}_1\otimes\hat{A}_2)(\psi\otimes\varphi)\doteq(\hat{A}_1\psi)\otimes(\hat{A}_2\varphi)$$

and on the expansion of a state in CONS as

$$(\hat{A}_1\otimes\hat{A}_2)\psi\doteq\sum_{i,j}c_{ij}\cdot(\hat{A}_1e_i)\otimes(\hat{A}_2f_j)$$

This definition implies that

$$((A_1\otimes B_1)+(A_2\otimes B_2))(\psi\otimes\varphi)=(A_1\otimes B_1)(\psi\otimes\varphi)+(A_2\otimes B_2)(\psi\otimes\varphi)$$
$$(A_1\otimes B_1)\cdot(A_2\otimes B_2)(\psi\otimes\varphi)=(A_1\cdot A_2)\psi\otimes(B_1\cdot B_2)\varphi$$

where the ”$\cdot$”-operation means multiplication of two operators.
EXAMPLE.

$$(\frac{\partial}{\partial x}\otimes\frac{\partial}{\partial y})(f\otimes g)(x,y)=(\frac{\partial}{\partial x}f\otimes\frac{\partial}{\partial y}g)(x,y)=\frac{\partial}{\partial x}f(x)\cdot\frac{\partial}{\partial y}g(y)$$

and

$$(\frac{\partial}{\partial x}\otimes\frac{\partial}{\partial y})(f)(x,y)=\frac{\partial}{\partial x}\frac{\partial}{\partial y}f(x,y)$$

We will use this operation systematically, anytime we add some new degree of freedom like other coordinates, other particles, other angular momenta and so on. Here we present examples related to the sum of two angular momenta – let it be the angular momenta of two different particles or the orbital and spin angular momenta of one single particle.

Example 1: The summing of the z-components of angular momenta in quantum mechanics. The operator of the total angular momentum of two subsystems, each having angular momentum quantum number J_1 and J_2 $(J_i = 0,\frac{1}{2},1,\frac{3}{2},....)$ is defined as

$$\hat{\mathbf{J}}\doteq\hat{\mathbf{J}}_1\otimes\mathbb{1}+\mathbb{1}\otimes\hat{\mathbf{J}}_2$$

or, component-wise

$$\begin{aligned}\hat{J}_x &= \hat{J}_{1,x}\otimes\mathbb{1}+\mathbb{1}\otimes\hat{J}_{2,x}\\ \hat{J}_y &= \hat{J}_{1,y}\otimes\mathbb{1}+\mathbb{1}\otimes\hat{J}_{2,y}\\ \hat{J}_z &= \hat{J}_{1,z}\otimes\mathbb{1}+\mathbb{1}\otimes\hat{J}_{2,z}\end{aligned}$$

In the literature one often use the abbreviated form

$$\hat{\mathbf{J}}\doteq\hat{\mathbf{J}}_1+\hat{\mathbf{J}}_2$$

and adopts the convention that the operator with the lower index 1 acts on the state residing on the left hand side, the operator with the lower index 2 acts on the state residing on the right hand side of the tensor product of states. For the specific case of the z-component we have

$$(J_{1z}+J_{2z})Y_{J_1}^{m_1}\otimes Y_{J_2}^{m_2}=(J_zY_{J_1}^{m_1})\otimes Y_{J_2}^{m_2}+Y_{J_1}^{m_1}\otimes J_zY_{J_2}^{m_2}$$

In Dirac writing

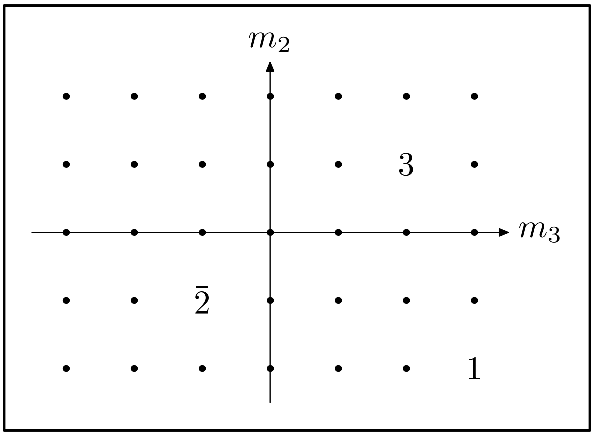
$$(J_{1z}+J_{2z})\,|\,J_1,m_1>|\,J_2,m_2>=(J_{1z}\,|\,J_1,m_1>)\,|\,J_2,m_2>+\,|\,J_1,m_1>(J_{2z}\,|\,J_2,m_2>)$$

THEOREM: The product states $Y_{J_1}^{m_1}\otimes Y_{J_2}^{m_2}$ with $m_1=J_1,\dots,-J_1$ and $m_2=J_2,\dots,-J_2$ are eigenfunctions of $J_{1,z}\otimes\mathbb{1}+\mathbb{1}\otimes J_{2,z}$ to the eigenvalues $\hbar\cdot(m_1+m_2)$

PROOF. One can use the computation rules $J_zY_l^m=\hbar mY_l^m$ and write

$$\begin{aligned}(J_{1,z}\otimes\mathbb{1}+\mathbb{1}\otimes J_{2,z})Y_{J_1}^{m_1}\otimes Y_{J_2}^{m_2} &= \underbrace{(J_{1,z}Y_{J_1}^{m_1})}_{\hbar m_1Y_{J_1}^{m_1}}\otimes Y_{J_2}^{m_2}+Y_{J_1}^{m_1}\otimes\underbrace{J_{2,z}Y_{J_2}^{m_2}}_{\hbar m_1Y_{J_2}^{m_1}}\\ &= \hbar(m_1+m_2)\cdot Y_{J_1}^{m_1}\otimes Y_{J_2}^{m_2}\end{aligned}$$

◇.



Cartan-Diagram for finding all possible eigenvalues of $\mathbf{3}_z\otimes\mathbb{1}+\mathbb{1}\otimes\mathbf{2}_z$*. It illustrates the use of Cartan diagrams for keeping track of the magnetic quantum numbers of a system consisting of subsystems.*

In order to facilitate the bookkeeping of these eigenvalues, Cartan has used a simple diagram summarizing them. Given a subsystem with angular momentum quantum number J_1 and a subsystem with angular momentum quantum number J_2, a Cartan lattice is defined by the set $\{J_1,J_1-1,\dots,-J_1\}\times\{J_2,\dots,-J_2\}$. Then all possible magnetic quantum numbers m_1+m_2 appear as a "weight" (m_1+m_2) associated to the lattice point (m_1,m_2). The Cartan diagram is applicable to both orbital and spin angular momenta.

Example 2: The summing of the quantum numbers of angular momenta in quantum mechanics. We have learned how to sum the z-components of the angular momenta, but the question we would like to answer now is: how does one sum the "length" of the vectors of the angular momenta? Remember that the angular momentum quantum number J is a measure of the "length" of the quantum mechanical angular momentum as it enters the eigenvalue of $\hat{\mathbf{J}}^2$ through the formula $\hbar^2 J(J+1)$! So, the question is: given an angular momentum with a certain "lenght" J_1 and a second angular momentum with a certain length J_2, what is the "length" (i.e. the quantum number) of the total angular momentum $\mathbf{J}_1 + \mathbf{J}_2$? In classical physics two vectors can be aligned parallel or antiparallel, or forming an angle so that the length of the total vector is some positive number in the interval $[J_1 + J_2, J_1 - J_2]$. Notice that merely reading out the Cartan weights is not enough to find the possible values of the total length. The situation is very similar to classical physics: two vectors can add to a different total length but have the same z-component. In quantum mechanics the "sum of lengths" is perfomed by an algorithm proposed by Cartan and is based on a particular strategy of "reading" the Cartan diagram.

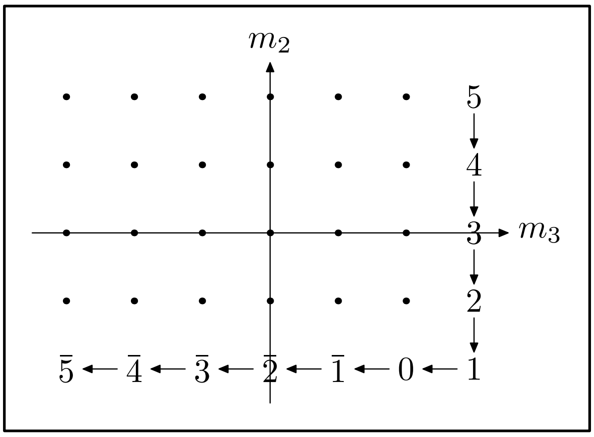


Illustration of the strategy to find the possible total angular momentum quantum numbers for the operator $\mathbf{3} \otimes \mathbb{1} + \mathbb{1} \otimes \mathbf{2}$.

- Pick up the lattice point with the highest Cartan weight – if you have just started, it is the lattice point with coordinate (J_1, J_2). Move vertically down until you have reached the end of the column and then move toward the left until you have reached the end of the line. In doing this, sample all Cartan weights you encounter during your walk. You will notice that you have sampled exactly $2 \cdot (J_1 + J_2) + 1$ Cartan weights. Think that they would perfectly fit as the $2(J_1 + J_2) + 1$ z-components of a total angular momentum quantum number $J_1 + J_2$! Conclude that you have found one of the possible values of the total angular momentum quantum number you were looking for!
- You have used Cartan weights – you may discard them. Go back to your diagram and pick up the highest Cartan weight remaining. If this is your second attempt, it will be $(J_1 - 1 + J_2)$. Start again your walk. At the end you will have used all Cartan weights corresponding to the total angular momentum quantum number $J_1 - 1 + J_2$!
- Discard the weights just found during the second walk, start again the walk and repeat it until you have used all Cartan weights! At the end you will have found all possible values for the angular momentum quantum numbers of the composite system (the "total angular momentum")

and you will notice that they obey a simple rule. If one defines the operation of sum of angular momentum quantum numbers with the symbol $\oplus$, the result is

$$J_1 \oplus J_2 = \begin{cases} J_1+J_2 \quad \text{or} \\ J_1+J_2-1 \quad \text{or} \\ \cdot \\ \cdot \\ \cdot \\ J_1-J_2 \end{cases}$$

This rule of summation is called as the Clebsch-Gordan series. This is the central result of this section and tells us how to sum angular momentum quantum numbers i.e. how to find the length of the sum of two angular momenta. It is apparent that summing angular momenta in quantum mechanics is an operation that is somewhat different than the arithmetic sum.

SUMMARY.
1. The operator $\hat{J}_{1z}+\hat{J}_{2z}$. Each product state

$$Y_{J_1}^{m_1} \otimes Y_{J_2}^{m_2}$$

is an eigenstate of the z-component of the total angular momentum operator: $\hat{J}_z = \hat{J}_{1,z} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{J}_{2,z}$. Its eigenvalues are given by

$$\hbar(m_1+m_2) \quad m_1 \in [J_1,...,-J_1] \quad m_2 \in [J_2,...,-J_2]$$

2. The operator $(\hat{\mathbf{J}}_1+\hat{\mathbf{J}}_2)^2$. The possible quantum numbers for the total angular momentum are given by

$$J_1+J_2, J_1+J_2-1, ..., J_1-J_2$$

These quantum numbers are <u>not</u> eigenvalues of any operator, but contain the necessary informations to give the eigenvalues of the operator of the square of the total angular momentum. This operator is defined, in a ”friendly” notation, as

$$(\hat{\mathbf{J}}_1+\hat{\mathbf{J}}_2)^2 = \left(\hat{J}_{x,1}+\hat{J}_{x,2}\right)^2 + \left(\hat{J}_{y,1}+\hat{J}_{y,2}\right)^2 + \left(\hat{J}_{z,1}+\hat{J}_{z,2}\right)^2$$

or, more precisely,

$$\left(\hat{J}_{1,x} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{J}_{2,x}\right)^2 + \left(\hat{J}_{1,y} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{J}_{2,y}\right)^2 + \left(\hat{J}_{1,z} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{J}_{2,z}\right)^2$$

This operator is difficult to handle but according to the rule of computations of angular momenta we do not need to solve explicitely its eigenvalue problem. We expect that in the subspace $\{Y_p^{m_p} \otimes Y_q^{m_q}\}$, the eigenvalues of this operators are given in terms of the quantum numbers $J_1+J_2, J_1+J_2-1, ..., J_1-J_2$ as follows:

a) The quantum number J_1+J_2 generates the eigenvalue

$$\hbar^2 \cdot (J_1+J_2) \cdot (J_1+J_2+1)$$

This eigenvalue is $2(J_1+J_2)+1$-times degenerate.
b) The quantum number J_1+J_2-1 generates the eigenvalue

$$\hbar^2 \cdot (J_1+J_2-1) \cdot (J_1+J_2-1+1)$$

This eigenvalue is $2(J_1+J_2-1)+1$-times degenerate.
c) The quantum number ... generate the eigenvalue

$$\hbar^2 \cdot (...) \cdot (...+1)$$

This eigenvalue is $2(...)+1$-times degenerate.
and so on.

A warning: the Clebsch-Gordan coefficients. We have found the eigenvalues of $(\hat{\mathbf{J}}_1+\hat{\mathbf{J}}_2)^2$ by using our knowledge on the quantum numbers, without doing one single bit of computation. We must, however, point out that the product functions

$$Y_{J_1}^{m_1} \otimes Y_{J_2}^{m_2}$$

are not necessarily eigenfunctions of the operator $(\hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2)^2$. It is a most intriguing and significant result of quantum mechanics that the eigenfunctions of the operator $(\hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2)^2$ in the subspace $\{Y_{J_1}^{m_1} \otimes Y_{J_2}^{m_2}\}$ are special linear combinations of these product functions – an example of **entanglement**! These special linear combinations are then eigenstates of both

$$\hat{J}_{z,1} + \hat{J}_{z,2} \quad \text{AND} \quad (\hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2)^2$$

simultaneously. These linear combinations build new basis states which play a crucial role in physics. The coefficients of the linear combinations are called CLEBSCH-GORDAN coefficients and are given in tables in the literature. We will introduce them at the right place when necessary.

6. Perturbation theory of discrete eigenvalues: the Ritz-method.

- As a starting point we consider an Hermitian operator $\hat{H}_0$ – typically the Hamilton operator of a system. The eigenvalues of $\hat{H}_0$ are supposed to be known and we concentrate on the section of the spectrum that contains discrete eigenvalues:
    - Eigenvalues $\{E_n^0\}$, n = 1,2, ...
    - Let E_n^0 be f-time degenerate:

      $$\{\varphi_{n,i}^0\}; \quad i = 1,...,f$$

      are the known eigenfunction to the eigenvalue E_n^0:

      $$\hat{H}_0 \varphi_{n,i}^0 = E_n^0 \varphi_{n,i}^0$$

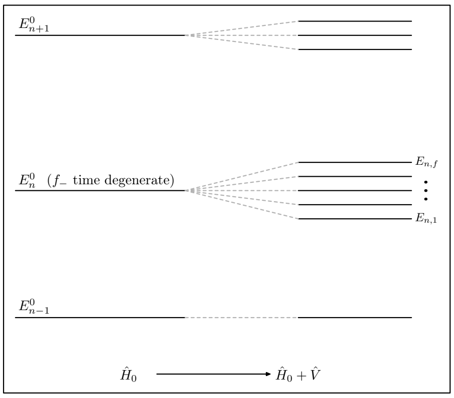
As an example, $\hat{H}_0$ is the Hamilton operator of the Hydrogen atom. We have solved this problem exactly and found a set of eigenvalues which are partly degenerate and the corresponding eigenfunctions.

- We add a small perturbation operator $\hat{V}$ to $\hat{H}_0$. The perturbation takes into account e.g. some further interaction that was excluded in the original operator $\hat{H}_0$ – e.g. with the aim of simplifying the solution of the eigenvalue problem of $\hat{H}_0$. An example of interactions which are not included in the original Hamilton operator of the hydrogen atom are those containing an external magnetic field. The complete operator of the system is now

$$\hat{H} \doteq \hat{H}_0 + \hat{V}$$

- The goal of perturbation theory is to specify formulas for the calculation of the eigenvalues and eigenfunctions of $\hat{H}$ starting from the known eigenvalues E_n^0 and the known wave functions $\varphi_{n,i}^0$ of the 'unperturbed' operator $\hat{H}_0$. A convenient situation for perturbation theory to be successful is that $\hat{V}$ represents a "small" perturbation of $\hat{H}_0$. This means that the change of the eigenvalue E_n^0 produced by the operator $\hat{V}$ is much smaller than the distance between E_n^0 and the neighbouring eigenvalues $E_{n\pm1}^0$

$$|E_n^0 - E_{n\pm1}^0| \gg |E_{n,1} - E_{n,f}|$$



• The method of Ritz (W. Ritz, Prof. ETH Zurich, 1897-1909) consists in establishing a finite dimensional Ritz space of trial wave functions within which the eigenvalue problem of $\hat{H}$ is solved. If we intend to find the small corrections to the eigenvalue E_n^0 produced by the perturbation operator $\hat{V}$, the Ritz space of choice is the eigenspace of the unperturbed operator $\hat{H}_0$ belonging to the eigenvalue E_n^0. Within this finite space, one solves the eigenvalue problem

$$\underbrace{\begin{pmatrix} \left(\varphi_{n,1}^0,\hat{V}\varphi_{n,1}^0\right) & \left(\varphi_{n,1}^0,\hat{V}\varphi_{n,2}^0\right) & . & . & . & \left(\varphi_{n,1}^0,\hat{V}\varphi_{n,f}^0\right) \\ \left(\varphi_{n,2}^0,\hat{V}\varphi_{n,1}^0\right) & \left(\varphi_{n,2}^0,\hat{V}\varphi_{n,2}^0\right) & . & . & . & \left(\varphi_{n,2}^0,\hat{V}\varphi_{n,f}^0\right) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \left(\varphi_{n,f}^0,\hat{V}\varphi_{n,1}^0\right) & \left(\varphi_{n,f}^0,\hat{V}\varphi_{n,2}^0\right) & . & . & . & \left(\varphi_{n,f}^0,\hat{V}\varphi_{n,f}^0\right) \end{pmatrix}}_{\text{Ritz matrix}} \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_f \end{pmatrix} = E_{n,i}^V \cdot \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_f \end{pmatrix}$$

The computed eigenvalues

$$E_n^0 + E_{n,i}^V$$

$i = 1, ...f$ are approximate eigenvalues of the complete operator $\hat{H}$ and the resulting linear combinations

$$\sum_{j=1}^{f} a_j^{(i)} \varphi_{n,j}^0$$

are the approximate eigenfunctions.

COMMENTS.

1. This is a linear homogeneous system of algebraic equations for the sought for coefficients $a_1, a_2, ...a_f$. It has a non-trivial solution only if the determinant of the $f \times f$ matrix

$$[\left(\varphi_{n,i}^0, \hat{V}\varphi_{n,j}^0\right)] - \mathbb{1}$$

vanishes.

2. The eigenvalue problem of the complete Hamiltonian operator $\hat{H}$ has been reduced, by the Ritz method, to the solution of the determinantal equation of a finite matrices.

3. The determinantal equation has certainly f real solutions, as $\hat{H}$ is supposed to be Hermitic. If all solutions are different, the original f-degeneracy is completed lifted and the original eigenvalue splits into f-different levels $E_{n,1}, ...E_{n,f}$. The eigenfunction to each of these levels is given by

$$\varphi_{n,i} = \sum_k a_k^{(i)} \varphi_{n,k}^0$$

In the basis $\{\varphi_{n,i}^0\}$ only $\hat{H}_0$ is diagonal. In the basis $\{\varphi_{n,i}\}$ both $\hat{H}_0$ and $\hat{H}$ are diagonal.

4. If some solutions of the determinantal equation coincide, then the degeneracy is only partially lifted and the eigenspaces to $E_{n,i}$ can be further degenerate.

5. If the original eigenvalue is non-degenerate ($f = 1$), then the perturbed eigenvalue trivially writes

$$E_n^0 + \left(\varphi_{n,1}^0, \hat{V}\varphi_{n,1}^0\right)$$

The perturbing Hamiltonian produces a shift of the original eigenvalue.

6. The perturbative computation of eigenvalues just illustrated provides the eigenvalues of $\hat{H}$ in 1. order approximation and the eigenvectors of $\hat{H}$ in 0th-order approximation. For the corrections

in the higher order approximations the wave function of the neighbouring eigenstates $E^0_{n\pm1}, E^0_{n\pm1}, \ldots$ must be included in the Ritz space.

7. In summary, to solve the eigenvalue problem of $\hat{H}$ in the vicinity of E^0_n, the recipe is ”work within the subspace spanned by the eigenvectors to E^0_n”.

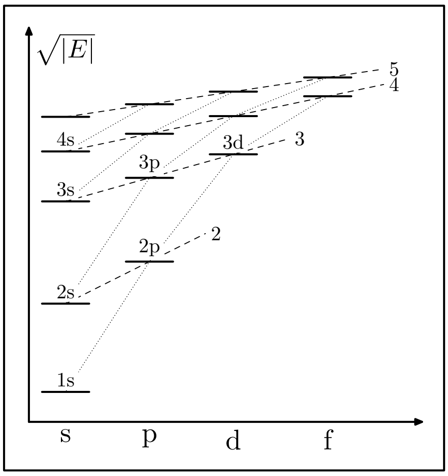
Part III.

Atomic and chemical physics

7. Energy levels, Pauli principle and the shell model of the atoms.

7.1. The energy levels of an atom.

For a system with atomic number Z there are Z electrons that interact via Coulomb interaction with the nucleus and with themselves. The problem of determining the energy levels of the electrons is a many-body one: the Coulomb repulsion among the electrons causes a complicated multi-particle problem whose solution can only be found approximately. There are many approximate methods for calculating the wave functions and energy states of atoms. These methods are described in advanced textbooks. The main result of these computational methods is the proof that in atoms the energy levels can be computed considering one single electron moving in the field of the nucleus AND a mean field of the remaining electrons (single-electron approximation, independent particle model). The mean field is found to have, to very good approximation, spherical symmetry. Accordingly, one can use the same terminology adopted for the hydrogen atom and label each energy level (so called " shell") by the quantum numbers n, l.



The energy levels of an electron in a spherically symmetric potential. Dashed lines join levels belonging to the same main quantum number. Dotted lines join levels with the same n_r-quantum number.

The one-electron approximation uses the operator

$$-\frac{\hbar^2}{2m}\triangle + \Phi_{\text{mean field}}(r)$$

to compute the energy levels of one electron in the mean field of the proton and all other electrons. The resulting energy levels are said to describe the configuration $(nl)^1$ and the operator used to generate them is known as the **configuration Hamiltonian**. In the next section, we will generate a

N-electron configuration by locating all the electrons forming an atom within these energy levels, starting from the lowest energy and according to a well defined recipe, called the ”shell” model. The final configuration will then be of the type $(1s)^x, (2s)^x, (2p)^x, ...$ until all Z-electrons are accomodated.

COMMENTS.
1. The most important result of solving the eigenvalue problem of the configuration operator within the independent model is that there is an energy difference between levels with the same n but different l. This means that the accidental degeneracy encountered in the hydrogen atom is lifted in more complex atoms, so that $E_{n,l-1} \neq E_{n,l-2} \neq ...E_{n,0}$. This difference can be understood on the basis of simple qualitative considerations, considering the effect of the field of the other electrons on the electron in question. To take this effect into account, one can use the wave functions of hydrogen-like atoms as a first approximation. The radial part of the wave function for states with a given orbital angular momentum (quantum number l) vanishes for $r \to 0$ as r^l. The electrons in s states can therefore come closer to the nucleus than the electrons with higher numbers l. Therefore, the full attraction of the nucleus on the electrons in s states is stronger than on the electrons in p, d, and f states. This causes an upward shit of the levels with higher l with respect to those with smaller l, see the figure. The states with given n, l quantum numbers form a shell (shell model of the electronic structure).

7.2. Pauli principle 1.0: The periodic table of the elements.

After estimating the energy levels of an electron in the Coulomb field of the nucleus and all other electrons using the one-electron approximation, we must specify how the Z electrons of an atom are distributed in these energy levels, i.e. how the levels are filled with electrons.
The electronic structure of the ground state of a multi-electron atom is built according to a ”modular construction principle” called the shell model. The most important rule that governs the filling of energy levels with electrons is the

PAULI PRINCIPLE (W. Pauli, 1925): ”In an orbital characterized by the quantum numbers (n, L, m_L), a maximum of two electrons may reside”.
Accordingly:
s states: Maximum 2 electrons
p states: Maximum 6 electrons
d states: Maximum 10 electrons
f states: maximum 14 electrons
etc.
1st shell $n = 1, s$: Maximum 2 electrons
2nd shell $n = 2, s + p$: Maximum 10 electrons
3rd shell $n = 3, s + p + d$: Maximum 18 electrons
etc.

The Pauli principle was formulated by Pauli in 1925 on the base of observing that certain atoms (with Z=2,8,18) were particularly stable, i.e. non-chemically reactive. Pauli proposed that an explanation of these ”magic numbers” could be found by filling Bohr quantized orbits with a given L and m_L with a maximum of 2 electrons and proposed that each electron in the state (L, m_L) assumes the opposite z-component of the newly found degree of freedom, the ”spin”. The filling of energy levels with electrons according to the Pauli principle produces one for each element characteristic electron configuration

$$(1s)^2(2s)^2(2p)^6(3s)^2......$$

indicating how many electrons occupy the np-shell. In order to find the ground state configuration, one fills progressively the energy levels, starting from the lowest one in energy an continuing with higher lying, until all electrons have found place. This produces the ground state configuration.
The shell structure of the electron states in the atoms, which follows from the quantum mechanical law of motion of the electrons, has in some sense been foreseen by the most important Russian chemist Mendeleyev in 1868, long before the emergence of quantum mechanics. Mendeleev discovered a periodic regularity of the chemical properties of elements which he then represented in the Periodic Table of the Elements along columns and lines. The periodic table of elements consists of horizontal rows – with our modern understanding the beginning of a new row is characterized by a transition to a higher number n – and vertical columns in which chemically similar elements are arranged – in our modern understanding these elements have a similar outer shell electron configuration. Not only the exact calculation of the energy levels of the hydrogen atom is a great triumph of the Schrödinger equation, but – in conjunction with the Pauli principle – also the explanation of Mendeleev's Table of Elements and the atomic Structure.

Let us now proceed with a brief construction of the periodic table as seen from the point of view of the shell model.

- Noble gases: $He, Ne, Ar, ...$ In the Helium atom the two electrons fill the first shell. Electronic configuration: $(1s)^2$. In the neon atom Ne_{10} two shells are completely occupied – configuration $(1s)^2(2s)^2(2p)^6$. Ar_{18} has three shells completely occupied – configuration $(1s)^2(2s)^2(2p)^6(3s)^6$ and so on. The atoms with closed electron shells form the noble gases. They have vanishing resulting orbital angular momentum and the total spin is zero. These atoms are very stable and do not undergo chemical bonds with other atoms, and they interact little with each other. They are, accordingly, arranged along the same column in the table. The particular chemical inactivity of these atoms must be assigned to their configuration, i.e. to having closed shells. This suggests a fundamental empirical principle that will have great application to chemistry: the "saturation principle" according to which a chemical compound builds to form as many as possible closed atomic shells. When a situation of as many as possible closed atomic shells is approached, the compound is "chemically stable " and thus the atoms stop undergoing further chemical reactions.

- Alkali metals: $Na, K, Rb, Cs, ...$ In each newly started shell, an electron is first inserted in a s state. All atoms with an electron outside of closed shells have similar chemical properties and belong to the alkali metals: $Li_3, Na_{11}, K_{19}, Cs_{55}, Fr_{87}$. They can also be arranged in the same column of a table.

- Transition metals: $V, Cr, Fe, Co, ...$ (3d), $Zr, Rh, Pd, ...$ (4d) $Ir, Pt, Au, ...$ (5d). When one gets to the main quantum number $n = 3$, one notices an exception: first $3s$ is filled, then $3p$ is filled further, but the $3d$ shell has an higher energy than the $4s$-shell, which is filled before the $3d$-shell (K: $(4s)^1$ then Ca with $(4s)^2$. Then the $3d$-shell starts being filled, whereby the configuration of the outer shell is kept fixed to the one of the Ca. Therefore, those elements which are formed by the filling of the $3d$-shell and have the same outer shell configuration belong to the same column of the periodic table (the second one) and must be separately ordered along an exceptional line – on a paper this is rendered by setting the d-transition metals together with the suitable element in the second column. The $3d$ elements, for instance, can be regarded as "offshoots" of the Ca (Ca is in the column of so-called "alkaline earth metals"). The d transition metals all have similar outer electron configurations s^2 (with a few exceptions), and can therefore be considered chemically equivalent.

- Lanthanide (= rare earths). The $4f$ shell is lower in energy than the $5d$ shell and will be filled accordingly. The 4f states are filled in after the element Lanthan La_{57}, which contains 54 electrons inside fully filled shells. The state of the three remaining electrons is determined by the configuration $(5d)^1,(6s)^2$. In the 14 elements, which are called rare earths or lanthanides, the $4f$ states are occupied. Since the electrons of the $4f$ states are accommodated in inner regions of the atom, the (spatially seen) outer shell remains almost unchanged that of the Lanthanum and the rare earths outer shell configuration is $(6s)^2$. The chemical properties of these elements are very similar to each other, and they belong to a place in the periodic table, the place of the Lanthanum.
- Actinides. In another group of elements, the electrons are incorporated into the atom by filling in the 5f shell without changing the configuration of the external electrons. The configuration of the outer electrons $(7s)^2$ is the same for all these elements.

7.3. Pauli principle 2.0: Multiplets

The single-electron approximation of the configuration Hamiltonian, supplemented by the Pauli principle in its most elementary form (the one of Pauli itself), provides electronic configurations where the many electrons of an atom find place. Such configurations, however, are generally degenerate, i.e. their eigenspace is built by a certain number of product functions. It is foreseable that the spin of each electron will sum up to produce a set of values S for the total spin angular momentum. Similarly, the orbital angular momenta will sum up to produce a set of values L for the total orbital angular momentum.

DEFINITION. The states characterized by the pair of values (L,S) hosted by a configuration are called ”multiplets”. The symbol for a multiplet is ^{2S+1}L.

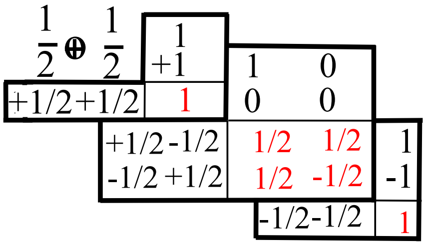
Organizing the basis states of a configuration in multiplet states is not just using a different set of basis states. The question is one of finding whether there is an interaction that breaks the degeneracy of the multiplets in a given configuration and which multiplet ends up to be the one with lower energy (the ground state) and which multiplets are the excited states, and what are the eigenfunctions in the subspaces belonging to the various multiplets. Which one of the multiplets is the ground state is an important physical question: depending on the values of L and S, the atom (and, ultimately, the matter made up of given atoms), will have different physical properties. For example, it will respond differently to a magnetic field. We need therefore to find the various multiplets and order them energetically.

• **The $(ns)^2$ configuration.** Let us consider, as an example of the situation, the simple configuration of two electrons in s-states, i.e. the $(ns)^2$-configuration. The four product basis states read (in the Dirac notation)

$$|n,0,\uparrow>|n,0,\uparrow> \quad |n,0,\uparrow>|n,0,\downarrow> \quad |n,0,\downarrow>|n,0,\uparrow> \quad |n,0,\downarrow>|n,0,\downarrow>$$

The ket on the left refer to the state of the first particle, the ket on the right to the state of the second particle. We now show that the configuration $(ns)^2$ sustains two multiplets. The total orbital angular momentum of the two electrons in s state can only be 0, i.e. $L=0$. To find the total spin quantum number, we must consider that the $(ns)^2$-configuration represents a composite system with two spin $\frac{1}{2}$-particles and, according to the rules of summation of angular momenta, the total spin quantum number of the composite system is either $S=0$ or $S=1$. The two possible multiplet-states are thus ^1S and ^3S. The multiplet ^1S is once degenerate: its basis function will be one linear combination of the four basis functions above. The multiplet ^1S is three times degenerate and its eigenspace will be determined by three linear combinations of the four basis functions above.

In order to find out more about the multiplets, we now search for those linear combinations of the basis states that are the eigenstates of $\mathbf{S}^2$. We know that these linear combinations are determined by the Clebsch-Gordan coefficients. These coefficients are often found in suitable Tables. We use the tables reported in M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018)). The table suitable for the sum of two $\frac{1}{2}$-spins is reproduced in the following figure.



$\frac{1}{2} \oplus \frac{1}{2}$	1	1	0	1
	+1	0	0	-1
+1/2 +1/2	1			
+1/2 -1/2		1/2	1/2	
-1/2 +1/2		1/2	-1/2	
-1/2 -1/2				1

Table for the Clebsch-Gordan coefficients of $\mathbf{S}^2 = 0$ *and* $\mathbf{S}^2 = 1$.

- How to read the table.
  - The top entries provide the quantum numbers of the sought-for eigenstates. In the specific case we expect the eigenstates to $\mathbf{S}^2 = 0$
$$|\,0,0>$$
and $\mathbf{S}^2 = 1$
$$|\,1,1> \quad |\,1,0> \quad |\,1,0>$$
For instance,
$$|\,1,0>$$
appears in the middle block as
$$\begin{matrix} 1 \\ 0 \end{matrix}$$
  - The left entries are the product basis states entering the eigenstates. For instance
$$|\!\uparrow>|\!\downarrow>\equiv|\,\frac{1}{2},-\frac{1}{2}>$$
appears in the second block on the left as
$$+\tfrac{1}{2} \quad -\tfrac{1}{2}$$
  - The numbers appearing at the intersection between vertical and horizontal lines are the actual Clebsch-Gordan coefficients (for simplicity of writing the sqrt-sign has been omitted in the tables). Accordingly, we read out from the table that
$$|\,1,0>=\sqrt{\frac{1}{2}}\,|\,\frac{1}{2},+\frac{1}{2}>+\sqrt{\frac{1}{2}}\,|\,\frac{1}{2},-\frac{1}{2}>$$

According to the table, the multiplet ${}^0S$ has the eigenstate
$$|\,n,L=0,m_L=0>\left(\sqrt{\frac{1}{2}}\,|\!\uparrow>\downarrow>-\sqrt{\frac{1}{2}}\,|\!\downarrow>\uparrow>\right)$$
The spin part of this $S=0$ state is referred in the literature as the "singlet" state. It is an eigenfunction of the total spin angular momentum $\mathbf{S}^2$ to the spin quantum number $S=0$ and also an eigenfunction

to the z-component $m=0$ of the total spin angular momentum.
The multiplet ${}^3S$ has the eigenstates

$$|\, n,L=0,m_L=0>\uparrow\uparrow> \quad |\, n,L=0,m_L=0>\left(\sqrt{\frac{1}{2}}\,|\uparrow>\downarrow>+\sqrt{\frac{1}{2}}\,|\downarrow>\uparrow>\right) \quad |\, n,L=0,m_L=0>|\downarrow\downarrow>$$

The spin part of these $S=1$ states are referred in the literature as the "triplet" states. They are eigenfunction of the total spin angular momentum $\mathbf{S}^2$ to the spin quantum number $S=1$ and also eigenfunctions to the z-components $m=1$, $m=0$ and and $m=-1$, respectively.

The rules of summing angular momenta in a composite system has produced, starting from the original four product functions, linear combinations for the subspaces of the respective multiplets. Besides of re-organizing the initial states, we have not been able to order energetically the multiplets. This is a quite common situation encountered when organizing the eigenspaces of any configuration into mutliplets. Something is still missing, that possibly favors one multiplet with respect to others.

7.3.1. Dirac formulation of the Pauli principle

The Pauli principle – a maximum of two electrons in an orbital – was suggested in 1925 on the base of empirical observations. A more precise and general formulation is due to Dirac (1926). It is based on the proper way of handling identical particles in quantum mechanics.
In order to illustrate the thinking of Dirac we take a system consisting of two electrons in the field of two protons – in fact the simplest two-electron model we can think of, i.e. the He-atom. The classical Hamilton functions writes

$$H(\mathbf{p}_1,\mathbf{p}_2,\mathbf{r}_1,\mathbf{r}_2)=\frac{\mathbf{p}_1^2}{2m}+\frac{\mathbf{p}_1^2}{2m}+\frac{e^2}{4\pi\varepsilon_0\,|\,\mathbf{r}_1-\mathbf{r}_2\,|}-\frac{2e^2}{4\pi\varepsilon_0\,|\,\mathbf{r}_1\,|}-\frac{2e^2}{4\pi\varepsilon_0\,|\,\mathbf{r}_2\,|}$$

This classical Hamilton function has an interesting symmetry: it is invariant with respect to an operation that changes $\mathbf{p}_1$ into $\mathbf{p}_2$, $\mathbf{p}_2$ into $\mathbf{p}_1$, $\mathbf{r}_1$ into $\mathbf{r}_2$ and $\mathbf{r}_2$ into $\mathbf{r}_1$. Such an operation is called "permutation" and written mathematically as

$$\Pi_{21}=\begin{pmatrix}\mathbf{p}_1 & \mathbf{p}_2\\ \mathbf{p}_2 & \mathbf{p}_1\end{pmatrix} \quad ; \quad \begin{pmatrix}\mathbf{r}_1 & \mathbf{r}_2\\ \mathbf{r}_2 & \mathbf{r}_1\end{pmatrix}$$

In fact, one can convince onself that the classical Hamilton function of N identical interacting particle is invariant with respect to any permutation of the particle coordinates and momenta – there are $N!$ of such permutations, according to the rule of mathematics. We know that in classical mechanics symmetries are very important: they produce conservation laws. For instance, the invariance of an Hamilton function with respect to rotation produces the conservation law for the angular momentum. Sadly, the invariance with respect to permutations does not have any relevant consequence in classical physics!
What about quantum mechanics? In quantum mechanics one must first translate the Hamilton function into an Hamilton operator. This is done by substituting $\mathbf{p}_1$ with $-i\hbar\nabla_1$ and $\mathbf{p}_2$ with $-i\hbar\nabla_2$, whereby $-i\hbar\nabla_i$ means that the partial derivatives are done with respect to the i-th variable. The Hamilton operator then reads

$$\hat{H}(\mathbf{r}_1,\mathbf{r}_2)=-\frac{\hbar^2}{2m}\cdot[\triangle_1+\triangle_2]+\frac{e^2}{4\pi\varepsilon_0\,|\,\mathbf{r}_1-\mathbf{r}_2\,|}-\frac{2e^2}{4\pi\varepsilon_0\,|\,\mathbf{r}_1\,|}-\frac{2e^2}{4\pi\varepsilon_0\,|\,\mathbf{r}_2\,|}$$

The Hamilton operator is invariant with respect to permutation of the two particle coordinates. It is realistic to assume that the Hamilton operator – and the other observable – of a system consisting of N identical particles are invariant with respect to permutations of the particle coordinates – including those coordinates that describe the spin. The question is: is there any practical consequence of this symmetry for the quantum mechanical behaviour of the N-particle systems?

The answer was found by Dirac. Let us see how it looks like for a two-particle system and then generalize it to N-particles. In quantum mechanics, a symmetry operation can always be associated with some operator acting onto the states. In the present case the symmetry operation is one of the two permutations

$$\Pi_{12} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

(this is the identical permutation) and

$$\Pi_{21} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

To account for these operations, Dirac introduced the permutation operators $\hat{P}_{12} \equiv \mathbb{1}$ and $\hat{P}_{21}$ acting in the space of states and defined as follows. Take a wave functions describing the state ψ of the two particles and containing the variables $\mathbf{r}_1, m_1, \mathbf{r}_2, m_2$. The variables $\mathbf{r}_i$ are continuous vectors in a three dimensional space, the variables m_i assume the discrete values $S, S-1, ..., -S$. For simplicity of writing, we join the two variables relating to the $i-th$- particle into one single variable ξ_i and the wave function is written as $\psi(\xi_1, \xi_2)$. Now define $\hat{P}_{ij}$ by let it act onto the wave function:

$$\hat{P}_{12}\psi(\xi_1,\xi_2) = \psi(\xi_1,\xi_2) \quad \hat{P}_{21}\psi(\xi_1,\xi_2) = \psi(\xi_2,\xi_1)$$

The operators just permute the arguments of every thing that resides on their right. Why did Dirac introduce these operators? Because it was aware of important mathematical consequences that are described by the following two theorems.

THEOREM 1: Provided $\hat{H}(\xi_1,\xi_2) = \hat{H}(\xi_2,\xi_1)$, the permutation operators commute with $\hat{H}$, i.e.

$$\left[\hat{P}_{i,j}, \hat{H}(\xi_1,\xi_2)\right] \doteq \hat{P}_{i,j}\hat{H}(\xi_1,\xi_2) - \hat{H}(\xi_1,\xi_2)\hat{P}_{i,j} = 0$$

or equivalently

$$\hat{P}^{-1}_{i,j}\hat{H}(\xi_1,\xi_2)\hat{P}_{i,j} = \hat{H}(\xi_1,\xi_2)$$

PROOF. We must show that $\hat{P}_{i,j}\hat{H}(\xi_1,\xi_2)\psi(\xi_1,\xi_2) = \hat{H}(\xi_1,\xi_2)\hat{P}_{i,j}\psi(\xi_1,\xi_2)$. We compute:

$$\hat{P}_{i,j}\hat{H}(\xi_1,\xi_2)\psi(\xi_1,\xi_2) = \hat{H}(\xi_2,\xi_1)\psi(\xi_2,\xi_1)$$

$$\hat{H}(\xi_1,\xi_2)\hat{P}_{i,j}\psi(\xi_1,\xi_2) = \hat{H}(\xi_1,\xi_2)\psi(\xi_2,\xi_1)$$

Provided $\hat{H}(\xi_1,\xi_2) = \hat{H}(\xi_2,\xi_1)$ the two expression of the right hand side are identical.
◇.

Why is the commutativity of two operators $\hat{A}$ and $\hat{B}$ so important? Suppose that one has solved the eigenvalue problem of the operator $\hat{A}$ and one has divided the Hilbert space into eigenspaces of the operator $\hat{A}$. Consider then the eigenspaces $\{\psi_i^1, \psi_i^2, ..., \psi_i^{l_i}\}$ to the eigenvalue A_i of $\hat{A}$. The commutativity allows to prove that $\{\psi_i^1, \psi_i^2, ..., \psi_i^{l_i}\}$ is also an eigenspace of $\hat{B}$! In other words:

THEOREM 2: commuting operators have common eigenspaces. This means that one can often find the eigenspaces of an operator $\hat{A}$ by studying other, perhaps simpler operators, provided they commute with $\hat{A}$.

PROOF: The proof of this theorem is as follows. Take one function ψ_i^j and operate first $\hat{B}$ and then $\hat{A}$ on it.

$$\hat{A}\hat{B}\psi_i^j = \hat{B}\hat{A}\psi_i^j = A_i\hat{B}\psi_i^j$$

The chain of equations proves that $\hat{B}\psi_i^j$ also belongs to the eigenspace $\{\psi_i^1, \psi_i^2, ..., \psi_i^{l_i}\}$ and that applying $\hat{B}$ to any eigenfunction ψ_i^j does not conduct us outside the eigenspace $\{\psi_i^1, \psi_i^2, ..., \psi_i^{l_i}\}$. The space $\{\psi_i^1, \psi_i^2, ..., \psi_i^{l_i}\}$ is also invariant with respect to $\hat{B}$.

$\diamond$.

Notice however that the function $\{\psi_i^1, \psi_i^2, ..., \psi_i^{l_i}\}$ need not to be eigenfunctions of $\hat{B}$ as well, but one can find l_i linear combinations of the functions $\{\psi_i^1, \psi_i^2, ..., \psi_i^{l_i}\}$ that diagonalize $\hat{B}$. At the end of this second diagonalization process one can find out that $\hat{B}$ might have more than one eigenvalue within the space $\{\psi_i^1, \psi_i^2, ..., \psi_i^{l_i}\}$ where $\hat{A}$ was degenerate! This second set of functions then diagonalize both $\hat{A}$ and $\hat{B}$, so that the theorem can be more sharply formulated as "Commuting operators have (after enough work is done) common eigenvectors".

The commuting of $\hat{P}$ and $\hat{H}$ means that we can find the eigenspaces of $\hat{H}$ by searching for the eigenspaces of $\hat{P}_{i,j}$. For this purpose, of particular interest is the operator P_{21}. The eigenvalue equation of this operator reads

$$\hat{P}_{2,1}\psi(\xi_1, \xi_2) = \lambda \cdot \psi(\xi_1, \xi_2)$$

with the sought for eigenvalue λ being some real number. We use the property

$$\hat{P}_{2,1}^2 = \mathbb{1}$$

To write

$$\begin{gathered}\hat{P}_{2,1}^2\psi(\xi_1, \xi_2) = \lambda^2 \cdot \psi(\xi_1, \xi_2) = \psi(\xi_1, \xi_2) \\ \rightarrow \lambda^2 = 1 \\ \rightarrow \lambda = \pm 1\end{gathered}$$

The eigenfunctions to the eigenvalue $+1$ must obey the equation

$$\hat{P}_{2,1}\psi(\xi_1, \xi_2) = \psi(\xi_2, \xi_1) = \psi(\xi_1, \xi_2)$$

In other words: the eigenfunctions to the eigenvalue $+1$ must be symmetric with respect to the permutation $(1,2) \rightarrow (2,1)$. The eigenfunctions to the eigenvalue -1 must obey the equation

$$\hat{P}_{2,1}\psi(\xi_1, \xi_2) = \psi(\xi_2, \xi_1) = -\psi(\xi_1, \xi_2)$$

In other words: the eigenfunctions to the eigenvalue -1 must be antisymmetric (change sign) with respect to the permutation $(1,2) \rightarrow (2,1)$. The symmetry of the Hamilton operator with respect to permutation of two identical particles has divided the possible eigenfunctions of the two-particle Hamilton operator into two sets: those symmetric and those antisymmetric with respect to a permutation of the particle coordinates. The spin coordinate must be interchanged as well, even if the Hamilton operator is not spin-dependent.

COMMENTS.

1. When N identical particles are involved, the above consideration are generalized as follows. The wave function describing the state ψ of the particles contain the variables $\mathbf{r}_1, m_1 \equiv \xi_1, \mathbf{r}_2, m_2 \equiv \xi_2, ..., \mathbf{r}_N, m_n \equiv \xi_N$: $\psi(\xi_1, \xi_2, ..., \xi_N)$. In general, by N particles we have the $N!$ permutations

$$\Pi = \begin{pmatrix} 1 & 2 & . & . & . & N \\ i_1 & i_2 & . & . & . & i_N \end{pmatrix}$$

and the $N!$ permutation operators

$$\hat{P}_\Pi\psi(\xi_1, \xi_2, ...\xi_N) = \psi(\xi_{i_1}, \xi_{i_2}, ..., \xi_{i_N})$$

The set of permutation builds a group and the rules of group theory allow to find a number of invariant subspaces into which the Hilbert space of a N-identical particle system is divided. Each "invariant subspace" has a well defined behaviour with respect to permutations. For instance, one invariant subspace is represented by symmetric wave functions – i.e. wave functions invariant with respect to all permutation operators – and a further invariant subspace is represented by antisymmetric wave functions – wave functions changing the sign if the permutation is uneven. There are, however, invariant subspaces where the wave functions have a more complicated behaviour.

2. A (yet) mysterious principle of Nature reduces greatly the number of subspaces allowed for identical particles: The Pauli exclusion principle, as formulated by Dirac, limits the subspaces that identical particles can occupy to those hosting either symmetric or antisymmetric wave functions. More precisely:

The states of a system of identical particles with integer spin (so-called bosons) lie in the subspace of the symmetric functions. The states of a system of identical particles with half-integer spin (so-called fermions) lie in the subspace of antisymmetric functions. Other symmetries or mixed symmetries do not occur in nature. On the base of this principle, electrons, protons and neutrons are fermions and assume antisymmetric wave functions. α-particles, or photons (spin 1) but also atoms at very low temperatures (see Bose-Einstein condensation) are bosons and assume symmetric wave functions.
3. The symmetry properties relate to the entire wave function, i.e the spin function of the entire system must be taken into account, even if $\hat{H}$ is spin-independent.

The $(ns)^2$-configuration

We are now ready to explore the $(ns)^2$-configuration and to classify the two multiplets ^3S and ^1S at the light of the requirement that the total wave function be antisymmetric with respect to permutation of the particle coordinates. We start with the multiplet ^1S and the associated spin singlet wave function (we omit the radial component for simplicity of writing):

$$Y_0^0(\Omega_1)\cdot Y_0^0(\Omega_2)\cdot\sqrt{\frac{1}{2}}\left[Y^+(m_1)\cdot Y^-(m_2)-Y^-(m_1)\cdot Y^+(m_2)\right]$$

When the particle coordinates are exchanged, the spin component becomes

$$\sqrt{\frac{1}{2}}\left[Y^+(m_2)\cdot Y^-(m_1)-Y^-(m_2)\cdot Y^+(m_1)\right]=-\sqrt{\frac{1}{2}}\left[Y^+(m_1)\cdot Y^-(m_2)-Y^-(m_1)\cdot Y^+(m_2)\right]$$

i.e. the spin component is antisymmetric with respect to coordinate exchange. The orbital component, instead, is symmetric. The total wave function is therefore antisymmetric with respect to particle exchange: the multiplet ^1S is allowed by the Pauli principle.
We turn now to the ^3S multiplet with wave functions

$$Y_0^0(\Omega_1)\cdot Y_0^0(\Omega_2)\cdot\left[Y^+(m_1)\cdot Y^+(m_2)\right]$$

$$Y_0^0(\Omega_1)\cdot Y_0^0(\Omega_2)\cdot\sqrt{\frac{1}{2}}\left[Y^+(m_1)\cdot Y^-(m_2)+Y^-(m_1)\cdot Y^+(m_2)\right]$$

$$Y_0^0(\Omega_1)\cdot Y_0^0(\Omega_2)\cdot\left[Y^-(m_1)\cdot Y^-(m_2)\right]$$

The three spin components are symmetric with respect to particle exchange. In order for the total wave function to be antisymmetric we need to construct an antisymmetric orbital component using one single orbital wave function. This is impossible: any attempt of antisymmetrizing the product wave functions of two electrons having individually exactly the same orbital wave function produces a vanishing wave function, i.e. given two electrons that occupy the same orbital state they can only form a spin singlet. The spin triplet is forbidden. This is a refined version of the original Pauli principle. For the specific configuration under investigation, the application of the Pauli principle has reduced the original degeneracy of the configuration and eliminated one multiplet.

COMMENTS.

1. We have found that all spin triplet states – corresponding to the various z-components of the total spin angular momentum – have exactly the same behaviour with respect to particle coordinate exchange. This is not an accident but it is protected by a famous theorem by I. Schur and H. Weyl:

THEOREM (Schur-Weyl duality)
All eigenfunctions to a given total spin S of an ensemble of spin $\frac{1}{2}$ particles have the same behaviour with respect to particle permutation. Accordingly, a given value of S provide a so-called ”Spin-Rasse”.

This theorem divides the termal scheme of an atom into well-separated termal schemes, each carrying a given value of S and each with well defined symmetry with respect to permutations. In addition, each ”Spin Rasse” determines the behaviour with respect to permutations of the orbital function allowed within the spin rasse, which has to be chosen so that the total wave function is finally antisymmetric. It might well happen that, in the process of building a spin rasse, some levels must be discarded from the termal scheme. For instance, the 1s-level appears in the $S=0$-termal scheme of the He atom as the ground state, while it is absent in the $S=1$ termal scheme.
2. Given a certain configuration and the corresponding multiplets, by virtue of the Pauli principle some of them are removed and the degeneracy of the configuration is reduced.

•**Slater determinants.** There is a technology that allows to construct antisymmetric wave functions starting from single particle wave functions. This technology is alternative to the one based on the Weyl theorem and has been invented by J.C. Slater. Consider e.g. the electronic configuration $(ns)^2$. The single particle states arising from the configuration Hamiltonian are (we omit the index n for simplicity of writing)

$$Y_{0,0}Y^+(\xi) \quad Y_{0,0}Y^-(\xi)$$

One can construct antisymmetrized product functions starting from these states by means of the so called Slater determinants:

$$\frac{1}{\sqrt{2}}\cdot\det\begin{pmatrix} Y_{0,0}Y^+(\xi_1) & Y_{0,0}Y^+(\xi_2) \\ Y_{0,0}Y^+(\xi_1) & Y_{0,0}Y^+(\xi_2)\end{pmatrix} = 0$$

$$\frac{1}{\sqrt{2}}\cdot\det\begin{pmatrix} Y_{0,0}Y^-(\xi_1) & Y_{0,0}Y^-(\xi_2) \\ Y_{0,0}Y^-(\xi_1) & Y_{0,0}Y^-(\xi_2)\end{pmatrix} = 0$$

$$\frac{1}{\sqrt{2}}\cdot\det\begin{pmatrix} Y_{0,0}Y^+(\xi_1) & Y_{0,0}Y^+(\xi_2) \\ Y_{0,0}Y^-(\xi_1) & Y_{0,0}Y^-(\xi_2)\end{pmatrix} = -\frac{1}{\sqrt{2}}\cdot\det\begin{pmatrix} Y_{0,0}Y^-(\xi_1) & Y_{0,0}Y^-(\xi_2) \\ Y_{0,0}Y^+(\xi_1) & Y_{0,0}Y^+(\xi_2)\end{pmatrix}$$

This last determinant provides the non-vanishing state

$$\frac{1}{\sqrt{2}}\cdot\left[Y_{0,0}Y^+(\xi_1)Y_{0,0}Y^-(\xi_2) - Y_{0,0}Y^-(\xi_1)Y_{0,0}Y^+(\xi_2)\right]$$

From this example one deduces the rule that Slater determinants involving identical states vanishes. This is in line with the original version of the Pauli principle that forbids two electrons having identical quantum numbers. The remaining two non-vanishing determinants produce the same linear independent antisymmetric product function that describes the spin singlet and represents a configuration where (sloppy) ”two electrons with opposite spin” can occupy identical orbital states.
The Slater determinant can be generalized to produce antisymmetric states involving N particles starting from M single particle states $\{\psi_{\alpha_i}\}$, $i = 1,...,M$:

$$\psi^-_{\alpha_1,\alpha_2,...,\alpha_N} = \frac{1}{\sqrt{N!}}\cdot\det\begin{pmatrix} \psi_{\alpha_1}(\xi_1) & \psi_{\alpha_1}(\xi_2) & . & . & . & \psi_{\alpha_1}(\xi_N) \\ \psi_{\alpha_2}(\xi_1) & \psi_{\alpha_2}(\xi_2) & . & . & . & \psi_{\alpha_2}(\xi_N) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \psi_{\alpha_N}(\xi_1) & \psi_{\alpha_N}(\xi_2) & . & . & . & \psi_{\alpha_N}(\xi_N)\end{pmatrix}$$

Taking into account that for the states used one must have $\alpha_i \neq \alpha_j$, there are

$$\begin{pmatrix} M \\ N \end{pmatrix}$$

such determinants. These states belong to multiplets within a given configuration that are allowed by Pauli principle. The technology of Slater determinants is the shortest path toward finding the antisymmetrized states within a given configuration. The Schur-Weyl theorem is the shortest path toward finding the multiplets within a given configuration.

The p^2-configuration

The Pauli principle has eliminated all but one product function in the $(s)^2$-configuration, leaving one Slater determinant and the corresponding multiplet – the spin singlet. We expect similar phenomena such as reduction of degeneracy and/or elimination of multiplets to occur in all configurations and they must be of course analyzed case by case in response to some experimental need. We conduct for the sake of illustrating these phenomena the analysis of the configuration of 2 electrons in the p-orbital. Without taking the Pauli principle into account, this configuration contains 36 product states. Using the technology of the Slater determinants we have 6 possible single-particle states and two electrons to place into them. This amounts to

$$\binom{6}{2} = 15$$

non-vanishing Slater determinants. The original degeneracy is therefore strongly reduced. The question is now about how to find the possible multiplets. Weyl theorem is more useful: it establishes that the multiplets will have either spin quantum number 0 (spin singlets, antisymmetric) or 1 (spin triplets, symmetric). Considering the orbital part of the wave function, we can build three possible total orbital quantum numbers from two $L = 1$ particles. In our terminology

$$1 \oplus 1 = 2, 1, 0$$

We find the corresponding states in the suitable Tables of Clebsch-Gordan coefficients (p. 564 of M. Tanabashi et al. (Particle Data Group), Phys. Rev. D **98**, 030001 (2018)), `https://journals.aps.org/prd/abstract/10.1103/PhysRevD.98.030001`). There we notice an important property:

- the $L = 2$ wave function is symmetric with respect to coordinate exchange, so that the $L = 2$ orbital wave function can only couple to the singlet wave function, producing 5 antisymmetric wave functions that belong to the quantum numbers $S = 0, L = 2$ and carry the multiplet symbol ^1D.
- The $L = 1$ wave functions, instead, are antisymmetric with respect to coordinate exchange and can only couple to the triplet spin states giving rise to 9 antisymmetric wave functions with multiplet symbol ^3P.
- The $L = 0$ wave function is also symmetric and produces the antisymmetric wave function with multiplet symbol ^1S.

The total number of antisymmetric states produced using Weyl theorem is, of course, 15, as expected from the counts of Slater determinants. In summary the configuration $(n, p)^2$ with energy $2 \cdot E_{np}$ hosts three multiplets:

$$^1D \qquad ^3P \qquad ^1S$$

We have a new situation with respect to the $(s)^2$ configuration: triplet states exist and provide possible non-vanishing antisymmetric states. The question now arises: among all the multiplets remaining, is there an interaction that – at least partially – removes the degeneracy and produces a ”multiplett splitting”? This is a key question toward determining which multiplet is the actual ground state of an atom and which ones are excited states. Notice that it makes a difference whether e.g. the ground state is a triplet or a singlet: in the first case the ensemble responds a la Zeeman (see next Chapter for the a treatement of the Zeeman effect) to a magnetic field, in the second the magnetic field just shifts the energy level.

7.4. The multiplet Hamiltonian: Estimate of multiplett splitting (Hund's rules).

We are now going to find out which interaction is responsible for the lifting of the multiplet degeneracy and find an estimate of the multiplet splitting energy by solving a simple model. We consider the $1s2s$-configuration of He. This is an excited state of He that provides two multiplets ^1S and ^3S which can be handled almost analytically, even using a quite complete two-body Hamiltonian. The order of magnitude we will obtain is applicable to more complex multiplets.
The configuration Hamiltonian reads (e^2 standing for $\frac{e^2}{4\pi\epsilon_0}$)

$$H_0(1,2) = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2}$$

The energy levels of $H_0(i)$ are $E_{n_i} = \frac{-Z^2e^2}{2\cdot a\cdot n_i^2}$, a being the Bohr radius. The eigenfunctions of E_{n_i} are

$$\varphi_{n_i,L_i,M_{L_i},m} = \frac{R_{n_i,L_i}}{r_i} \cdot Y_{L_i,M_{L_i}}(\vartheta_i, \varphi_i) \otimes Y^m$$

• **The ground state.** The ground state of H_0 corresponds to the state in which both electrons are in a $1s$-orbital. The configuration $(1s)^2$ has the energy

$$E_{(1s)^2} = 2\cdot E_{1s} = 2\cdot -\frac{Z^2e^2}{2\cdot a} = -4\cdot 13.6eV$$

The wave functions read

$$\varphi_{1s}(1)\varphi_{1s}(2) \otimes Y^{m'} \otimes Y^m = \frac{1}{\pi}(\frac{Z}{a})^3 e^{-\frac{Z}{a}(r_1+r_2)} \otimes Y^{m'} \otimes Y^m$$

$m, m = \pm$. After anti-symmetrization, only one wave function remains in the ground state, namely

$$\varphi_{1s}(1)\varphi_{1s}(2) \otimes \chi_s$$

χ_s being the singlet spin state.
• **The $(1s)(2s)$ configuration.** In order to get more than one multiplet in He one has to get to an excited state configuration. In atoms, a many-multiplet structure can happen in the ground state configuration, as worked out for the $(np)^2$ configuration. For energy estimates, however, the He-excited state is perfectly representative.
The configuration $(1s)^1(2s)^1$ is degenerate with respect to the configuration Hamiltonian. The anti-symmetrized wave functions read:

$$\frac{1}{\sqrt{2}}\cdot(\varphi_{1s}(\mathbf{r}_1)\varphi_{2s}(\mathbf{r}_2) - \varphi_{1s}(\mathbf{r}_2)\varphi_{2s}(\mathbf{r}_1)) \otimes \chi_t^1$$
$$\frac{1}{\sqrt{2}}\cdot(\varphi_{1s}(\mathbf{r}_1)\varphi_{2s}(\mathbf{r}_2) - \varphi_{1s}(\mathbf{r}_2)\varphi_{2s}(\mathbf{r}_1)) \otimes \chi_t^0$$
$$\frac{1}{\sqrt{2}}\cdot(\varphi_{1s}(\mathbf{r}_1)\varphi_{2s}(\mathbf{r}_2) - \varphi_{1s}(\mathbf{r}_2)\varphi_{2s}(\mathbf{r}_1)) \otimes \chi_t^{-1}$$
$$\frac{1}{\sqrt{2}}\cdot(\varphi_{1s}(\mathbf{r}_1)\varphi_{2s}(\mathbf{r}_2) + \varphi_{1s}(\mathbf{r}_2)\varphi_{2s}(\mathbf{r}_1)) \otimes \chi_s^0$$

χ_t are the triplet spin states. The energy of the $(1s)^1(2s)^1$ configuration is $E_{1s} + E_{2s}$.
• **Perturbation of the ground state by the multiplet Hamiltonian.** We now perturb the configuration Hamiltonian with the Coulomb repulsion between the two electrons

$$V_{1,2} = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{e^2}{r_{12}}$$

and obtain the so called "multiplet Hamiltonian"

$$H_0 + V_{1,2}$$

The Ritz matrix the this Hamiltonian in the subspace

$$\varphi_{1s}(1)\varphi_{1s}(2)\chi_s$$

reads

$$\left(E_0+Q_G\right)$$

with

$$Q_G=\int dV(1)dV(2)\varphi^2_{1s}(1)\frac{e^2}{r_{12}}\varphi^2_{1s}(2)$$

In order to explicitly compute this last integral, one develops $1/r_{12}$ in spherical harmonics:

$$\begin{aligned}\frac{1}{\mid \mathbf{r}_1-\mathbf{r}_2\mid} &= \frac{4\pi}{r_1}\sum_{l,m}\frac{1}{2l+1}(\frac{r_2}{r_1})^l\cdot Y^*_{l,m}(\vartheta_1,\varphi_1)Y_{l,m}(\vartheta_2,\varphi_2)\Longleftrightarrow r_1>r_2\\ \frac{1}{\mid \mathbf{r}_1-\mathbf{r}_2\mid} &= \frac{4\pi}{r_2}\sum_{l,m}\frac{1}{2l+1}(\frac{r_1}{r_2})^l\cdot Y^*_{l,m}(\vartheta_1,\varphi_1)Y_{l,m}(\vartheta_2,\varphi_2)\Longleftrightarrow r_1<r_2\end{aligned}$$

Inserting this development in Q_G and using the orthogonality of spherical harmonics we obtain

$$Q_G=\frac{4e^2}{\pi}(\frac{Z}{a})^6\int_0^\infty dr_1r_1^2\cdot e^{-\frac{2Zr_1}{a}}[\frac{1}{r_1}\int_0^{r_1}dr_2r_2^2e^{-\frac{2Zr_2}{a}}+\int_{r_1}^\infty dr_2r_2e^{-\frac{2Zr_2}{a}}]$$

Partial integration leads to $Q_G=\frac{5Ze^2}{8a}$ and positive. Notice that the correction Q_G is of the same order of magnitude as E_0.
The energy of the ground state configuration is modified by the Coulomb repulsion between electrons to

$$E_G=-\frac{Ze^2}{a}(Z-\frac{5}{8})$$

The only multiplet possible in the ground state is ${}^1S$. The net magnetic moment is vanishing: $S=1$ is prohibited in the ground state by the Pauli principle.

• **Perturbation of the ${}^1S$ and ^{3} multiplets by the multiplet Hamiltonian.** The Ritz matrix for the perturbed Hamiltonian $H_0+V_{1,2}$ in the subspace of the four ${}^1S$ and ^{3} multiplets wave functions of the $1s2s$-configuration reads:

$$\begin{pmatrix} E_{1s}+E_{2s}+Q_{1s2s}+J_{1s2s} & 0 & 0 & 0\\ 0 & E_{1s}+E_{2s}+Q_{1s2s}-J_{1s2s} & 0 & 0\\ 0 & 0 & E_{1s}+E_{2s}+Q_{1s2s}-J_{1s2s} & 0\\ 0 & 0 & 0 & E_{1s}+E_{2s}+Q_{1s2s}-J_{1s2s}\end{pmatrix}$$

with

$$Q_{(1s)(2s)}=\int dV\frac{\mid\varphi_{1s}(\mathbf{r}_1)\mid^2\cdot e^2\cdot\mid\varphi_{1s}(\mathbf{r}_2)\mid^2}{r_{12}}$$

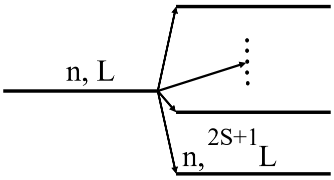
being the electrostatic interaction between the two charges densities $\mid\phi_{1s}\mid^2,\mid\phi_{2s}\mid^2$. $Q_{(1s)(2s)}$ is also expected from classical electrostatics and produces, as for the ground state energy, a shift of the energy of the $(1s)(2s)$ configuration. The integral J

$$J_{(1s)(2s)}=\int dV\frac{\varphi_{1s}(\mathbf{r}_1)^*\varphi^*_{2s}(\mathbf{r}_2)\cdot e^2\cdot\varphi_{1s}(\mathbf{r}_2)\varphi_{2s}(\mathbf{r}_1)}{r_{12}}$$

is called the **exchange integral** contribution, arising from the correlation of the two electron as a consequence of anti-symmetrizing the wave functions according to the Pauli principle and it is of purely quantum mechanical origin. $J_{(1s)(2s)}$ vanishes in the classical limit. $J_{(1s)(2s)}$ *produces a removal of the degeneracy between the multiplets and a split of the configuration energy level*. The strength of the splitting – the strength of the parameter $J_{(1s)(2s)}$ – is typically of the order of *eV* ($11605k_B$), i.e. the same order of magnitude as $Q_{(1s)(2s)}$. This is not surprising, as both $Q_{(1s)(2s)}$ and $J_{(1s)(2s)}$ originate from the same interaction. The sign is positive, so that the triplet has a lower energy than the singlet. There is an intuitive explanation for this positivity. The Coulomb energy is large when the two electrons are closer to each other. In a triplet spin state the antisymmetric orbital wave functions the two electrons takes care that the two electrons are as far as possible from each other, an this reduces the Coulomb repulsion with respect to the symmetric orbital wave function, where the two electrons are, on the average, allowed to be closer to each other.

COMMENTS.

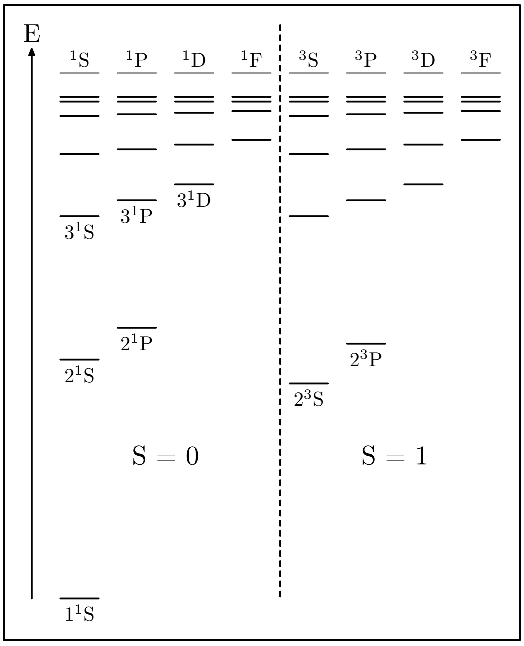
- This quantitative result is a-posteriori justification of a famous ”rule” obtained by F. Hund (1925) and Russel and Saunders (Astrophysics Journal, **61**, 38, 1925) on the basis of empirical atomic spectra observations ($1st$ Hund rule): ” *Multiplet splitting occurs in such a way that the triplett spin state has the lowest energy”. Equivalently: Provided there is sufficient degeneracy that non-equivalent orbital wave functions can be constructed, the configuration realizing the lowest energy state corresponds to a state of maximum spin number*. In other words: if the orbital states involved have different quantum numbers, the filling of the electronic states with parallel spins produces the lowest energy electronic configuration. Thus, provided orbitally degenerate states exist, the triplet state is energetically favoured. If degeneracy is absent, then the singlet state is energetically favoured, because of the Pauli principle.
- By these considerations, we have found an important result that guides the formation of a finite total spin in an atomic system: orbital degeneracy, i.e. a configuration carrying more than one multiplet – is required.
- There is a rule ($2nd$ Hund rule) that further produces splitting of multiplets, should any degeneracy be left behind after applying the first Hund rule: among the remaining multiplets, the one that has the lowest energy is the one with highest L. Again, this is because of the intuitive argument that the Coulomb energy is reduced for high values of L.
- The level scheme resulting from the multiplet Hamiltonian is summarized schematically in the next figure.



The energy level of the configuration Hamiltonian (left) is splitted into multiplets (right) by the Coulomb repulsion between electrons.

- Regarding the physics of He-atoms: there are two distinct termal schemes for He, consisting of para (singlet) - and ortho (triplet) states. Disregarding the spin-orbit interaction, transitions involving the emission or absorption of light between the triplet and the singlet states are forbidden (due to the orthogonality of the spin functions). Because of this, singlet and triplet states of the helium atom are independent. If a He atom reaches the lowest triplet state $(1s)(2s)$, it will remain in that state for longer times (months), even though it is an excited state. Because of this long lifetime, such states are called metastable. He atoms in singlet and triplet states can therefore be considered as two different types of atoms. Para-helium has no magnetic moment and forms a diamagnetic gas. Ortho-helium, on the other hand, forms a paramagnetic gas.

The levels in the t-states split due to the spin orbit interaction. Therefore, the spectral lines of Para-helium are simple lines. In contrast, the spectral lines of ortho-helium consist of closely adjacent lines (have a fine structure, see next chapter), and even show a hyperfine structure (see next chapter).



The term scheme of parahelium (left) and orthohelium (right). Particularly clear is the triplet-singlet splitting that systematically favours triplet states over singlet states energetically. Triplet states have a fine structure as well (too small to appear in the energy scale used in the figure.

7.5. The Heisenberg-Van Vleck-Dirac operator

One can formally obtain the $t-s$ splitting by caricaturing the exchange interaction (which is actually acting in the orbital space) with an effective spin Hamiltonian: the Heisenberg-Dirac-Van Vleck operator. Dirac defined an operator acting in spin space

$$H_{Spin} = (E_{1s} + E_{2s} + Q_{1s2s}) \cdot \mathbb{1} - J_{1s2s} \cdot P_{12}$$

with the exchange operator P_{12}

$$P_{12} Y^{\pm} Y^{\pm} = Y^{\pm} Y^{\pm}$$

$$P_{12} Y^{\pm} Y^{\mp} = Y^{\mp} Y^{\pm}$$

exchanging the z-components of the two particles. The eigenvalues of the operator H_{spin} when restricted to the $1s2s$ eigenspace, are identical with the eigenvalues of the physical operator. A useful way of writing P_{12} is

$$P_{12} = \frac{\mathbb{1} + \sigma_1 \otimes \sigma_2}{2}$$

or, setting $\mathbf{S} = \frac{1}{2}\sigma$

$$P_{12} = \frac{1}{2}[\mathbb{1} + 4 \cdot \mathbf{S}_1 \otimes \mathbf{S}_2]$$

where σ are the two-by-two Pauli matrices and the product $\sigma_i \otimes \sigma_j$ must be taken as a Kronecker product of matrices. The spin Hamiltonian simulating the original Hamiltonian acting in the orbital space reads

$$H_{Spin} = (E_{1s} + E_{2s} + Q_{1s2s}) \cdot \mathbb{1} - J_{1s2s} \cdot [\frac{1}{2}\mathbb{1} + 2\mathbf{S}_1 \otimes \mathbf{S}_2]$$

The non-diagonal part, generalized to any spin and some exchange parameter J

$$-2 \cdot J \cdot \mathbf{S}_1 \otimes \mathbf{S}_2$$

is also called the Heisenberg exchange (vector) operator. A simplified version of it

$$-2 \cdot J \cdot S_1^z \otimes S_2^z$$

involving only the z-component of the spin operator is called the "Ising model". Versions of the Heisenberg and Ising models where **S** is a classical vector are also used. In physical science, but also in any other branch of natural science and even in humanities, the Heisenberg and Ising models, in particular in their generalization to many degrees of freedom

$$-2 \cdot J \cdot \sum_{i \neq j} \mathbf{S}_i \otimes \mathbf{S}_j$$

respectively

$$-2 \cdot J \cdot \sum_{i \neq j} S_{z_i} \otimes S_{z_j}$$

have a widespread use. They represent the simplest way of capturing and describing two-body interactions – such as those appearing in a community of fishes or in the traffic of cars moving within restricted roads – and lead to non-trivial complexity.

8. Fine structure, Zeeman effect and hyperfine splitting.

8.1. The fine structure of atomic levels.

After having sorted out multiplets according to the first and second Hund rule, we might still have some degeneracy left. Let us go back to the $(np)^2$ configuration. There we found that the configuration $(n,p)^2$ with energy $2 \cdot E_{np}$ hosts three multiplets: $^1D, ^3P, ^1S$. Let us order them according to the Hund rules. The multiplet with the lowest energy is certainly the ^3P multiplet. ^1D lies higher, followed by ^1S (2nd Hund rule). With respect to the total angular momentum quantum number J, ^1D and ^1S are not degenerate, as they carry one single quantum number J, namely $J = 2$ (^1D) and $J = 0$ (^1S). The ^3P-multiplet, instead, is degenerate with respect to the quantum numbers $J = 2, 1, 0$ for the total angular momentum. The question is now: is there any interaction that is capable of removing this degeneracy and assigning to multiplets with different J different energies? Experimentally, this question was answered by the observation of the Na-doublet and was called by A. Sommerfeld the **fine structure** of the energy levels.

The origin of the fine structure is an interaction that corrects the multiplet Hamiltonian: the so called **spin-orbit interaction**. It is the result of taking the theory of relativity into account when formulating the Hamilton operator and describes a coupling between the spin degree of freedom and the orbital degree of freedom. The derivation of the spin-orbit coupling operator from the Dirac equation is beyond the scope of this lecture. We refer to advanced literature on the subject(Masud Mansuripur,"Spin-orbit coupling in the hydrogen atom, the Thomas precession, and the exact solution of Dirac's equation", Proc. SPIE 11090, Spintronics XII, 110901X (16 September 2019); `https://doi.org/10.1117/12.2529885`. The spin-orbit coupling operator reads[1]

$$H_{LS} = A_{LS} \cdot \underbrace{[\hat{L}_x \otimes \hat{S}_x + \hat{L}_y \otimes \hat{S}_y + \hat{L}_z \otimes \hat{S}_z]}_{\hat{\mathbf{L}} \otimes \hat{\mathbf{S}}}$$

In the literature, A_{LS} has the unit of an energy and the operators $\hat{L}$ and $\hat{S}$ are taken to be dimensionsless. We will use this convention henceforth. This operator can be rewritten using the identity

$$(\hat{\mathbf{L}} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathbf{S}})^2 = \hat{\mathbf{L}}^2 \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathbf{S}}^2 + 2 \cdot \hat{\mathbf{L}} \otimes \hat{\mathbf{S}}$$

as

$$H_{LS} = A_{LS} \cdot \frac{(\hat{\mathbf{L}} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathbf{S}})^2 - \hat{\mathbf{L}}^2 \otimes \mathbb{1} - \mathbb{1} \otimes \hat{\mathbf{S}}^2}{2}$$

This last expression contains explicitly the square of the operator of the total angular momentum

$$\hat{\mathbf{J}}^2 \doteq (\hat{\mathbf{L}} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathbf{S}})^2$$

Our having introduced an interaction that explicitly contains the square of the operator of the total angular momentum is important to understand why this interaction is able to split the J-degenerate levels within a given multiplet. Within the set of basis states of a given multiplet ^{2S+1}L

$$\{| L, m_L > | S, m_S >\}$$

---

[1] The spin-orbit coupling constant has a dependence on the radial variable. However, as spectroscopic observations relate to the average over the radial charge density of the coupling constant $A_{LS}(r)$, we prefer to use the averaged quantity A_{LS} to write the effective spin-orbit coupling operator.

- the operator

$$\mathbf{L}^2 \otimes \mathbb{1}$$

is diagonal and has the $(2L+1)\cdot(2S+1)$ degenerate eigenvalue

$$L(L+1)$$

- the operator

$$\mathbb{1} \otimes \mathbf{S}^2$$

is diagonal and has the $(2L+1)\cdot(2S+1)$ degenerate eigenvalue

$$S(S+1)$$

- the operator

$$(\mathbf{L}\otimes \mathbb{1} + \mathbb{1} \otimes \mathbf{S})^2$$

is not diagonal but we know its eigenvalues – without solving the determinantal problem: Why do we know the eigenvalues? Because the operator

$$(\mathbf{L}\otimes \mathbb{1} + \mathbb{1} \otimes \mathbf{S})^2$$

is the square of the total angular momentum, and we have learned how to compute all the possible angular momentum quantum numbers of a composite system with orbital quantum number L and spin quantum number S:

$$L \oplus S = L+S \quad L+S-1 \quad ..... \quad L-S$$

Knowing the possible values of the total angular momentum quantum number we have a rule on how to comute the eigenvalues of the square of the vector of the total orbital angular momentum:

$$L+S \to (L+S)\cdot(L+S+1) \quad ..... \quad L-S \to (L-S)(L-S+1)$$

Accordingly, we have solved the eigenvalue problem of the spin-orbit coupling operator within the space of product function without diagonalizing it explicitly:

$$E_J = A_{LS} \cdot \frac{J(J+1)-L(L+1)-S(S+1)}{2}$$

$J = L+S, ..., L-S$. This result confirms that

$$E_J - E_{J-1} = A_{LS} \cdot J$$

This rule was found empirically by Lande in 1923 (Lande- interval rule).

COMMENTS.

1. The spectroscopic symbols that are used to identify the various levels are:

a) the symbol indicating the configuration

b) the symbol indicating the possible multiplets – i.e. values of L and S that can be realized in this configuration: ${}^{(2S+1)}L$.

c) A given multiplet can be split into **terms** with different J, which are indicated by the spectroscopic symbol

$${}^{2S+1}L_J$$

Terms have different energy because of the spin-orbit splitting.

2. In the case of one electron in a p level, for instance, we have two possible terms:

$${}^2P_{J=\frac{3}{2}} \quad \text{and} \quad {}^2P_{J=\frac{1}{2}}$$

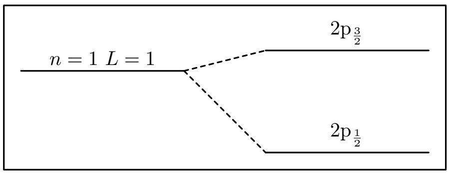


Figure 8.1.: *Spin-orbit split energy levels.*

3. the ^3P multiplet splits into three terms:

$$^3P_2$$
$$^3P_1$$
$$^3P_0$$

• 4. **Third Hund's rule.** The order of the levels is ruled by the $3rd$ Hund rule, which originates from the eigenvalue expression for the terms we have just derived. The third rule states that the " *total angular momentum quantum number J minimizing the energy is $|L+S|$ if the shell is more than half full, $|L-S|$ if the shell is less than half full.*"

5. The strength of the coupling constant A_{SL} can be computed precisely for H-similar atoms. With

$$A_{LS}(r) = \frac{Z \cdot \hbar^2 \cdot l^2}{4\pi\epsilon_0 \cdot 2m^2c^2r^3}$$

we find

$$A_{SL} = \int_0^\infty r^2 dr f_{nL}^2(r) \cdot \frac{Z \cdot \hbar^2 e^2}{4\pi\epsilon_0 2m^2c^2r^3} = \frac{\alpha^2 \cdot Z^4}{2n^3(l+1)(l+\frac{1}{2}) \cdot l} \cdot \overbrace{Ha}^{Hartree}$$

Plugging in the universal constant

$$\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}$$

(the so called fine-structure constant), we estimate $A_{SL} \sim 10^{-4} - 10^{-5} Ha$. For the $n=2$ level of the H-atom ($Z=1$) the energy difference between the terms

$$^2P_{\frac{3}{2}} \quad ^2P_{\frac{1}{2}}$$

is $\frac{\alpha^2}{32}$ Ha. This very small splitting justifies a posteriori our choosing as Ritz space the space belonging to the multiplet ^{2S+1}L, without involving other configurations.

6. The absolute value of the fine structure splitting decreases with increasing main quantum number but increases rapidly with the fourth power in Z. For heavy atoms, corrections to term energies are therefore relevant and produce the failure of 1st order perturbation theory and, accordingly, a remixing of the energy levels that produces some exceptions in the periodic table of the elements.

8.2. An atom in the external magnetic field.

8.2.1. Excursus in classical magnetostatic

For treating the problem of an electron in a magnetic field we need to construct the suitable Hamilton operator. The proper strategy could be to translate the Hamilton function of a charged particle in an electric and magnetic field into a corresponding Hamilton operator. At disposal from classical electrodynamics are the relations

$$\mathbf{E} = -\nabla\phi(\mathbf{r},t) - \frac{\partial \mathbf{A}}{\partial t}$$

and

$$\mathbf{B} = \nabla \times \mathbf{A}$$

which relate the electric and magnetic field to the scalar potential $\phi(\mathbf{r},t)$ and the vector potential $\mathbf{A}(\mathbf{r},t)$ and the Newton equation of motion

$$m\ddot{\mathbf{r}} = q \cdot \mathbf{E} + q(\dot{\mathbf{r}} \times \mathbf{B})$$

In classical physics, one can set the Hamilton function as Ansatz and prove that the Hamilton equations of motion lead to the Newton equation of motion.

THEOREM: The Hamilton function for a charge q in an electromagnetic field writes

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + q\phi$$

PROOF:
We recall: The Hamilton equations for the variables $\mathbf{r}$ and $\mathbf{p}$ write

$$\dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{p}}; \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{r}}$$

With $\mathbf{r} \doteq (x_1, x_2, x_3)$ and $\mathbf{p} \doteq (p_1, p_2, p_3)$ we set up to compute

$$m\ddot{x}_i = m\frac{d}{dt}\frac{\partial H}{\partial p_i} = \dot{p}_i - q\dot{A}_i$$

Now:

$$\begin{aligned} \dot{p}_i &= \frac{\partial H}{\partial x_i} = \frac{1}{m}(\mathbf{p} - q\mathbf{A}) \cdot q \cdot \frac{\partial \mathbf{A}}{\partial x_i} - q\frac{\partial \phi}{\partial x_i} \\ &= \sum_j \dot{x}_j q \frac{\partial A_j}{\partial x_i} - q\frac{\partial \phi}{\partial x_i} \end{aligned}$$

and

$$\dot{A}_i = -q\frac{\partial A_i}{\partial t} - q\sum_j \frac{\partial A_i}{\partial x_j}\dot{x}_j$$

so that

$$m\ddot{x}_i = -q\left(\frac{\partial \phi}{\partial x_i} + \frac{\partial A_i}{\partial t}\right) + q\sum_j \dot{x}_j\left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j}\right)$$

which is the sought for Newton equation

$$m\frac{d^2\mathbf{r}}{dt^2} = q \cdot \mathbf{E} + q(\dot{\mathbf{r}} \times \mathbf{B})$$

◇

Before constructing the Hamilton operator from the Hamilton function with the correspondence principle, let us obtain a formally more useful version of the classical Hamilton function. We restrict ourselves to a uniform magnetic field $\mathbf{B}$ so that we can write

$$\mathbf{A} = -\frac{1}{2}\mathbf{r} \times \mathbf{B}$$

Inserting in the Hamilton function we obtain

$$\begin{aligned} H(\mathbf{p}, \mathbf{r}, \mathbf{A}) &= \frac{1}{2m}(\mathbf{p} + \frac{q}{2}\mathbf{r} \times \mathbf{B})^2 + q\phi = \frac{\mathbf{p}^2}{2m} + q\Phi + \underbrace{\frac{q}{2m} \cdot \mathbf{p} \cdot (\mathbf{r} \times \mathbf{B})}_{-\frac{q}{2m}(\mathbf{r}\times\mathbf{p})\cdot\mathbf{B}} + \frac{q^2}{8m}(\mathbf{r} \times \mathbf{B})^2 \\ &= +\frac{\mathbf{p}^2}{2m} + q\Phi \to (a) \\ &\quad -\frac{q}{2m}(\mathbf{L} \cdot \mathbf{B}) \to (b) \\ &\quad +\frac{q^2}{8m}(\mathbf{r} \times \mathbf{B})^2 \to (c) \end{aligned}$$

Strength estimate.

- The term labeled (a) produces the Hamilton operator encountered in the original work by Schrödinger on the Hydrogen atom. The electronic part a is of the order of eV, compute e.g.

$$\left(\psi_{ns}, \left(-\frac{\hbar^2}{2m}\triangle + q\phi\right)\psi_{ns}\right)$$

for hydrogen like wave functions to find that

$$| E_{2s} - E_{1s} | \approx 10.2 \ \ eV$$

• Take e.g. **B** along the positive z-direction and compute e.g.

$$\left(\psi_{2p,m=1}, (\frac{|e|}{2m} \cdot L_z \cdot B_z) \psi_{2p,m=1} \right) = \frac{|e| \hbar}{2m} \cdot B_z$$

to find that the characteristic strength of b, assuming an orbital angular momentum $\hbar$, is of the order of

$$10^{-4} \frac{eV \cdot B}{T} \quad B \quad \text{in units of Tesla}$$

(b) is therefore $<< (a)$, even when magnetic field strengths of few Tesla – which are typical for laboratory using superconducting coils – are used.

• Let **B** be along $+z$ and compute (c) for a $1s$-state of Hydrogen

$$\left(\psi_{1s}, (\frac{e^2}{8m}(x^2 + y^2) B_z^2) \psi_{1s} \right) = \left(\psi_{1s}, (\frac{e^2}{12m}(x^2 + y^2 + z^2) B_z^2) \, \psi_{1s} \right)$$

to find that (c) amounts to about

$$10^{-11} eV \cdot (\frac{B}{T})^2$$

c contributes a positive energy which is, for typical Tesla fields, much smaller than the strength of both a and b.

COMMENTS.

1. In classical physics the angular momentum is not quantized and a most peculiar phenomenon appears: the angular momentum can be small enough for b and c to be of the same strength. When averaging with the suitable Gibbs probability at any finite temperature over all possible classical orbits to compute the energy, b and c cancel out exactly, rendering the average energy independent of B. This exact compensation between b and c is the content of Bohr-van Leuwen theorem. In quantum mechanics, where L is quantized, our estimate shows that, unless the atomic configuration has exactly $L = 0$ (it can happen, e.g. for noble gas elements), the term (b) always dominates in strength with respect to (c) and is therefore the only one needed to be taken into account for explaining e.g. spectroscopic observations of atoms in a magnetic field. From now on we will therefore neglect c and work with b as a small perturbation of a, respectively b as a small perturbation of the fine structure levels.

2. In classical physics,

$$q \cdot \mathbf{L} \doteq \mu$$

is the so called vector of the **magnetic moment** and (b) is also written as

$$-\mu \cdot \mathbf{B}$$

This terminology is maintained when the operators are translated to quantum mechanics.

8.2.2. The Zeeman effect.

The translation of the coupling Hamiltonian between a charge q and magnetic field **B** to quantum mechanics must take into account that quantum mechanical particles have an internal degree of freedom – the spin. A first attempt to build the quantum mechanical operator is to use the correspondence principle and write

$$\frac{e}{2m} \cdot \mathbf{L} \cdot \mathbf{B} \quad \rightarrow \quad \frac{e}{2m} \cdot (\hat{\mathbf{L}} + \hat{\mathbf{S}}) \cdot \mathbf{B}$$

for a set of electrons, each with charge $q = -e$ and mass m, total orbital angular momentum $\hat{\mathbf{L}}$ and total spin angular momentum operator $\hat{\mathbf{S}}$. This translation was long thought to be correct but could not explain spectroscopic observations of atoms. In 1928, Dirac equation proved this translation to be wrong: the correct **Zeeman operator**[2] for an ensemble of electrons in a magnetic field is

$$H_Z(\mathbf{B}) = \underbrace{\frac{e\cdot\hbar}{2m}}_{\mu_B} \cdot (g_L \hat{\mathbf{L}} + g_S \cdot \hat{\mathbf{S}}) \cdot \mathbf{B}$$

$g_L = 1$ and $g_S = 2$ are the so called g-factors for the orbital respectively spin degrees of freedom. The value of g_S is a yet non understood consequence of the relativistic Dirac equation[3]. The coupling constant

$$\frac{e\cdot\hbar}{2m} \doteq \mu_B = 9.274009994(57) \times 10^{-24} \frac{J}{T}$$

is the so called Bohr magneton. With this coupling constant, the operators $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ become dimensionless[4] We want to find the eigenvalues of the operator H_Z in the subspace of the eigenfunctions of the fine structure level ${}^{2S+1}L_J$

$$\{|J, m_J>\}$$

i.e. we want to diagonalize the matrix

$$< J, m'_J \,|\, \hat{H}_Z \,|\, J, m_J >$$

whereby, without loss of generality, we can set $\mathbf{B} = (0,0,B)$ and

$$H_Z = \mu_B \cdot B \cdot (g_L \hat{L}_z + g_S \hat{S}_z)$$

In this subspace the operators $\mathbf{J}^2$ and J_z are diagonal. What about $\hat{L}_z$ and $\hat{S}_z$? There is a systematic property of the eigenspaces $\{|J, m_J>\}$ that is the consequence of a celebrated theorem of mathematical physics: the Wigner-Eckart-Koster theorem. In virtue of this theorem, one can show that the matrix elements of H_Z are independent on J, i.e both L_z and S_z are diagonal matrices of the type

$$\hat{S}_z = \kappa_S \cdot \hat{J}_z$$

and

$$\hat{L}_z = \kappa_L \cdot \hat{J}_z$$

κ_S and κ_L being some parameters dependent on L, S, J but – and this is most important – independent on m_J. κ_L is determined from

$$\kappa_L \hat{\mathbf{J}}^2 = \hat{\mathbf{L}} \cdot \hat{\mathbf{J}} = \frac{1}{2}(\hat{\mathbf{J}}^2 + \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2)$$

κ_S follows from

$$\kappa_S \hat{\mathbf{J}}^2 = \hat{\mathbf{J}}^2 - \kappa_L \hat{\mathbf{J}}^2$$

Inserting these results, the Zeeman matrix writes

$$g_{LSJ} \cdot \mu_B \cdot B \cdot \begin{pmatrix} J & 0 & 0 & . & . & 0 \\ 0 & J-1 & 0 & . & . & 0 \\ 0 & 0 & J-2 & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & 0 & -J \end{pmatrix}$$

g_{LSJ} is the Lande factor:

$$g_{LSJ} = \frac{1}{2} \frac{g_L(J(J+1) + L(L+1) - S(S+1)) + g_S(J(J+1) + S(S+1) - L(L+1))}{J(J+1)}$$

---

[2] Pieter Zeeman, Nobel prize 1902)

[3] It is not even exactly "2": Quantum electrodynamic provides the value 2.00231930436182 – and this is one of the physical quantities which have been measured with the highest precision. We will continue using "2".

[4] Using the operator

$$\hat{\mu} \doteq -\mu_B \cdot (g_L \hat{\mathbf{L}} + g_S \cdot \hat{\mathbf{S}})$$

the Zeeman operator writes

$$H_Z = -\hat{\mu} \cdot \mathbf{B}$$

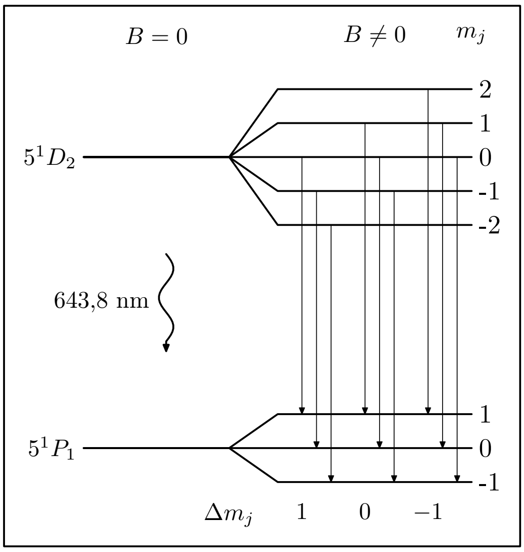
or, inserting $g_L = 1$ and $g_S = 2$

$$g_{LSJ} = 1 + \frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}$$

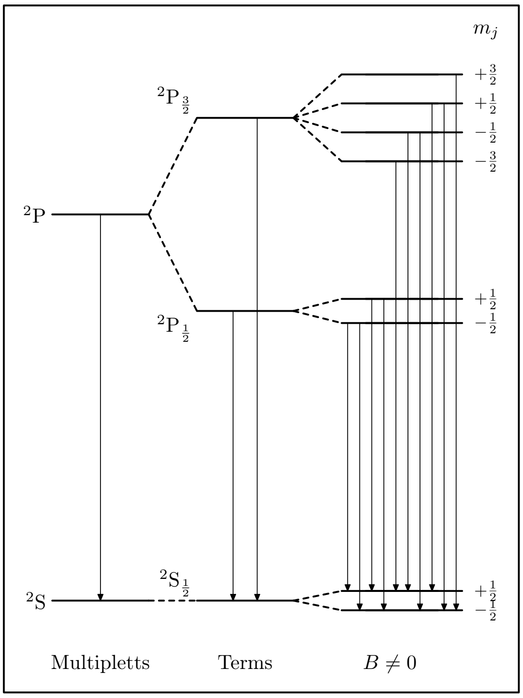
It is a diagional matrix, so that its eigenvalues can be read out from the diagonal.
COMMENTS.
1. When a fine structure level L,S,J is embedded into a magnetic field, its degeneracy is completely lifted and the level spits into $2J+1$ sub-levels
2. The distance between the levels is

$$\mu_B \cdot B \cdot g_{LSJ}$$

3. The distance depends on the quantum numbers L,S,J – this phenomenon is known as the anomalous Zeeman effect. It is to be compared with the "normal" Zeeman effect that foresees an universal distance between sub-levels of just $\mu_B \cdot B$. The non-universality is a consequence of the gyromagnetic factor of the spin angular momentum being 2 and not 1. The observation of a non-universal distance between sub-level in a magnetic field was a big mystery (an unexplained "anomaly") that could only be solved when the Dirac equation appeared, which contained explicitely $g_S = 2$



Schematic diagram for the Zeeman effect in the red spectral line of a Cd atom. It corresponds to the transition
$5^1D_2(J=2, L=2, S=0) \rightarrow 5^1P_1(J=1, L=1, S=0)$
The levels have $S=0$ and the Zeeman effect is the "normal" one: both lines are split by the magnetic field into equidistant sublevels. Using left-circularly polarized light ($\Delta m_J = 1$), right circularly polarized light ($\Delta m_J = -1$) or linearly polarized light ($\Delta m_J = 0$) one observes only three optical transitions.



Schematic diagram for the Zeeman effect in Na atom. It corresponds to the transitions
$^2P_{\frac{3}{2}}(J=\frac{3}{2},L=1,S=\frac{1}{2}) \rightarrow ^2S_{\frac{1}{2}}(J=\frac{1}{2},L=0,S=\frac{1}{2})$
and
$^2P_{\frac{1}{2}}(J=\frac{1}{2},L=1,S=\frac{1}{2}) \rightarrow ^2S_{\frac{1}{2}}(J=\frac{1}{2},L=0,S=\frac{1}{2})$
The anomalous Zeeman effect is observed with a total of 10 *optical transitions with left-circularly polarized light (*$\Delta m_J = 1$*), right circularly polarized light (*$\Delta m_J = -1$*) and linearly polarized light (*$\Delta m_J = 0$*)*

8.3. * The hyperfine splitting of s-states

When calculating the fine-structure and Zeeman corrections, we considered the field of the atomic nucleus to be a spherically symmetric electric field. The nucleus of the hydrogen atom and many other atoms have a magnetic moment. The interaction between the magnetic moment of the electron and that of the nucleus further remove the degeneracy of the atom's energy levels. The nuclear magnetic moments are about 10^3 times smaller than the magnetic moment of the electron. The splitting of the levels due to the nuclear magnetic moment will therefore be about 10^3 times smaller than the splitting due to the electronic magnetic moment and the fine structure splitting. It is therefore called hyperfine splitting. The measurement of the hyperfine structure of an atom is e.g. a method for determining experimentally the spins and the magnetic moments of the atomic nuclei. We want to estimate the HF-splitting of the s-states of an electron in an atom. We can look at the atomic nucleus as a point like particle carrying a spin $\hbar\cdot\mathbf{I}$. This spin produces an atomic magnetic moment and a corresponding Dirac-delta like magnetization vector:

$$\frac{|\,q\,|\,\hbar}{2m_K}\cdot g_K\cdot\mathbf{I}\cdot\delta(\mathbf{r})$$

q is the charge of the nucleus and for a proton $g_K\cong 5.586$. This magnetization produces a magnetic field $\mathbf{B}$ that interacts "a-la-Zeeman" with the magnetic moment of the electron. In order to compute $\mathbf{B}$ we use a relation of classical magnetostatics:

$$\mathbf{A}(\mathbf{r})=\frac{\mu_0}{4\pi}\cdot\nabla\times\int dV'\frac{\overbrace{\mathbf{M}(\mathbf{r}')}^{\frac{|e|\hbar}{2m_K}\cdot g_K\cdot\mathbf{I}\cdot\delta(\mathbf{r})}}{|\mathbf{r}-\mathbf{r}'|}=\beta\cdot\nabla\times\frac{\mathbf{I}}{|\mathbf{r}|}$$

with β summarizing the various constants:

$$\beta\doteq\frac{\mu_0}{4\pi}\frac{|\,e\,|\,\hbar}{2m_K}\cdot g_K$$

The magnetic field $\mathbf{B}$ is now computed from the vector potential $\mathbf{A}$, using some relations from vector analysis:

$$\begin{aligned}\mathbf{B} &= \nabla\times\mathbf{A}(\mathbf{r})=\beta\cdot\nabla\times\left(\nabla\times\frac{\mathbf{I}}{|\mathbf{r}|}\right)\\ &= -\beta\cdot\mathbf{I}\cdot\triangle\frac{1}{|\mathbf{r}|}+\nabla\left(\nabla\cdot\frac{\mathbf{I}}{|\mathbf{r}|}\right)\\ &= \beta\cdot 4\pi\delta(\mathbf{r})\cdot\mathbf{I}+\beta\cdot\frac{3(\mathbf{I}\mathbf{r})\mathbf{r}-r^2\mathbf{I}}{|\mathbf{r}|^5}-\beta\cdot\frac{4\pi}{3}\delta(\mathbf{r})\cdot\mathbf{I}\end{aligned}$$

When s-states are involved, only that part of $\mathbf{B}$ which contains Dirac-delta like functions is relevant. The spherical symmetry of the s-wave function produces the vanishing of the remaining part, when the spatial averaging in the matrix element is performed. This is, however, not true for p and d states, where the usual dipole-like magnetic field term also contributes to the matrix element. The Hamilton operator relevant for the HF-interaction in s-levels reads:

$$H_{HF}=-\vec{\mu}_S\cdot\mathbf{B}=\frac{8\pi}{3}\cdot\beta\cdot\mu_B\cdot\sigma\cdot\mathbf{I}\cdot\delta(\mathbf{r})$$

where precisely speaking, the scalar product $\sigma\cdot\mathbf{I}$ means

$$\sigma_x\otimes I_x+\sigma_y\otimes I_y+\sigma_z\otimes I_z$$

The Ritz space for solving the eigenvalue problem is given by the set of basis states

$$\{\psi_{n,l=0}(r)\otimes Y^{m_s}\otimes Y_I^{m_I};\quad m_s=\pm;\quad m_I=I,\ldots,-I\}$$

The matrix elements in this subspaces amounts to

$$\left(\psi_{n,l=0}(r)\otimes Y^{m_s'}\otimes Y_I^{m_I'},(\frac{8\pi}{3}\cdot\beta\cdot\mu_B\cdot\sigma\cdot\mathbf{I}\cdot\delta(\mathbf{r}))\psi_{n,l=0}(r)\otimes Y^{m_s}\cdot Y_I^{m_I}\right)$$

$$= \frac{8\pi}{3} \cdot \beta \cdot \mu_B \underbrace{\left(\psi_{n,l=0}(r), \delta(\mathbf{r})\psi_{n,l=0}(r)\right)}_{|\psi_{n,l=0}(0)|^2} \cdot \left(Y^{m_s'} \cdot Y_I^{m_I'}, \sigma \cdot \mathbf{I} Y^{m_s} \cdot Y_I^{m_I} \right)$$

The eigenvalue problem of the $2(2I+1) \times 2(2I+1)$ matrix

$$[\left(Y^{m_s'} \cdot Y_I^{m_I'}, \sigma \otimes \mathbf{I}\, Y^{m_s} \cdot Y_I^{m_I} \right)]$$

is solved by the usual algebraic "trick". Let us introduce the operator

$$\mathbf{F} \doteq \mathbb{1} \otimes \mathbf{I} + \frac{1}{2} \sigma \otimes \mathbb{1}$$

of the total angular momentum of the electron and the nucleus. From

$$\mathbf{F}^2 = \mathbb{1} \otimes \mathbf{I}^2 + \frac{1}{4} \sigma^2 \otimes \mathbb{1} + \sigma \cdot \mathbf{I}$$

we obtain

$$\sigma \otimes \mathbf{I} = \mathbf{F}^2 - \mathbb{1} \otimes \mathbf{I}^2 - \frac{1}{4} \sigma^2 \otimes \mathbb{1}$$

The possible values for

$$\frac{1}{4}\sigma^2 \quad \text{are} \quad \frac{1}{2} \cdot (\frac{1}{2} + 1)$$

The eigenvalues of

$$\mathbf{I}^2 \quad \text{are} \quad I(I+1)$$

The possible eigenvalues for

$$\mathbf{F}^2 \quad \text{are} \quad (I \pm \frac{1}{2}) \cdot \left(I \pm \frac{1}{2} + 1 \right)$$

Using these results we can compute the sought for eigenvalues for the operator $\sigma \otimes \mathbf{I}$:

$$F(F+1) - I(I+1) - \frac{3}{4} = I \quad F(F+1) - I(I+1) - \frac{3}{4} = -I - 1$$

COMMENTS.

1. The s-states split, in virtue of the hyperfine interaction, into two states with the energy

$$E_I = E_{n,s} + \frac{8\pi}{3} \cdot \beta \cdot \mu_B |\psi_{n,l=0}(0)|^2 \cdot I \quad E_{-I-1} = E_{n,s} - \frac{8\pi}{3} \cdot \beta \cdot \mu_B |\psi_{n,l=0}(0)|^2 \cdot (I+1)$$

2. The HF-splitting amounts to

$$E_I - E_{-I-1} = \frac{8\pi}{3} \cdot \beta \cdot \mu_B |\psi_{n,l=0}(0)|^2 \cdot (2I+1)$$

Inserting $\beta \doteq \frac{\mu_0}{4\pi} \frac{|e|\hbar}{2m_K} \cdot g_K$ and $|\psi_{n,l=0}(0)|^2 = \frac{Z^3}{\pi a^3 n^3}$ produces

$$E_I - E_{-I-1} = \frac{\mu_0}{4\pi} \frac{|e|\hbar}{4m_K m} \cdot g_K \cdot \frac{Z^3}{\pi a^3 n^3} \cdot (2I+1)$$

3. For $n = 2$ in the Hydrogen atom one gets

$$\frac{E_I - E_{-I-1}}{2\pi\hbar} \simeq 1420\text{MHz}$$

This splitting corresponds to a wavelength $\lambda \cong 21cm$ for $n = 2$ in a H atom. This characteristic radiation is e.g. emitted by hydrogen clouds in the galaxies and is used in astronomy to determine the position and concentration of hydrogen masses in space.

9. Quantum chemistry

The main tasks of quantum chemistry is to explain how atoms can undergo a chemical bond to form molecules and – ultimately – solids. Specific features of the chemical bond, like the bond between identical atoms, its saturation properties – an hydrogen atom binds with exactly one single other hydrogen atom, or a carbon atom can bind exactly four hydrogen atoms – and the particular chemical inactivity of noble gas atoms – together with a large amount of experimental observations – point to the fact that the chemical bond is related to the electronic configuration of the outer shells of the atomic components. Quantum mechanics provides a quantitative explanation of the simple rule that govern chemical bonds: the atoms undergo a chemical bond with the aim of filling their partially filled orbitals in such a way that an electronic configuration is built which resembles as close as possible to the configuration of a noble gas atom. Although usually the bonding mechanisms are mixed, four types of chemical bonds can be distinguished roughly in molecular chemistry:

- **Ionic or electrostatic bonding**. In this type of bonding, one observes complete transfer of electrons from the anion to the cation, when by this transfer the atomic electron configurations of both components of the bond are rendered similar to those of some noble gases. This ionic chemical bond is of course very strong between atoms of the first columns (which build the cations) and atoms of the 7th column (which build the anions), and the compound built are typically NaCl, KI, CsCl,etc. The attractive interaction that seals the chemical bond is provided by the Coulomb interaction between the ions. Also for most of semiconductor compounds there is a slight transfer of charge.

- **Covalent (or homo-polar) bonding:** Only few elements are well-defined metals and non-metal atoms such that the electron transfer can really provide the main mechanism of bonding. The largest part of the chemical bonds cannot be realized by electron transfer, in particular when identical atoms are involved, like the stable molecules H_2, O_2, N_2 etc. The chemical bond without noticeable transfer of electrons from an atom to another is called as homo-polar or covalent bond. The nature of this bonding is purely quantum mechanical and originate from the possibility of electrons to ”tunnel” from one atom the the other. The covalent bonding is essentially localized and the electrons are shared by the neighbors centers. The covalent bonding is dependent on the electron orbitals, consequently the resulting molecular orbitals have well defined directions in space. This bonding is of course the bond of choice between to identical atoms and appears therefore also when crystals of the same element are built (Si,Ge,...).

- **Delocalized (aromatic) bonding**. The resulting molecular orbitals overlap strongly so that the chemical bonding cannot be localized in space (although it has some modulation). In chemistry, this bonding type appears e.g. in the benzene molecule C_6H_6. Some of the valence electrons of the C undergo covalent bond strongly localized in between the various atoms. The p_z orbitals perpendicular to the plane overlap strongly and produce delocalized π-orbitals. In this orbital, one electron is shared by all C-atoms in the benzene ring.

- **Van der Waals's bonding**. The physical reason is the polarization of electron shells of the atoms and resulting dipole-dipole interaction. This bonding is typical for inert gas atoms (Ar, Xe, Cr, molecular crystals) and very weak. It leads to the formation of solids of noble atoms at very low temperature.

9.1. Covalent bonding: the H_2^+ ion

Several essential elements of the covalent bond can be deduced by computing the simplest molecule one can think of: H_2^+ ion. Because of the great mathematical difficulties in the treatment of many particle problems, one must resort to many approximations when computing the chemical bond. One of them is called the Born-Oppenheimer approximation. It consists of freezing the motion of the nuclei, which are then deprived of their kinetic energy. The only degree of freedom left behind is their distances, which enter the problem as a parameter to be adjusted for minimizing the total energy. Regarding the formation of the chemical bond, we point out that on one side there is the unbounded state, consisting of a proton at infinite distance from an hydrogen atom. The total ground state energy of the separated components amounts to the energy of the 1*s* energy level of hydrogen. On the other side, the proton is moved to a finite distance R from the hydrogen atom. We are now seeking some equilibrium distance R_0, at which the total energy of the system is lower than the energy of the unbounded state. The total Hamilton operator writes, in the Born-Oppenheimer approximation, and ignoring spin orbit coupling

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{r_A} - \frac{e^2}{r_B} + \frac{e^2}{R}$$

r_A and r_B indicate the distance between the electron and the proton A respectively B. e^2 is a shorthand symbol for $\frac{e^2}{4\pi\epsilon_0}$. We only have one electron, so we do not need to take the Pauli principle or the Hund rules into account. The Ritz space appropriate to find the solution of the eigenvalue problem of this Hamilton operator is taken to consist of the two hydrogen like wave functions $\psi_{A(B)}(\mathbf{r})$

$$(\pi a^3)^{-1/2} e^{-r_{A(B)}/a}$$

$a = \frac{\hbar^2}{me^2}$ being the Bohr radius. These wave functions describe the two possible ground states for $R \to \infty$. There are of course an infinite number of possible basis states, but a linear combination of the atomic ground state wave functions is a more likely candidate for the ground state of the molecule. In addition, by restricting the number of wave functions, one obtain a mathematically treatable problem. If we want a more accurate computation, we should take also the excited states of the hydrogen atom into account, but the size of the matrix increases. Although physically meaningful, such atomic orbitals have a computational drawback: they are not, in general, orthogonal, and the so-called overlap integral

$$(\psi_A, \psi_B) = (\psi_B, \psi_A) = \int dV \psi_A(\mathbf{r}) \cdot \psi_B(\mathbf{r}) = \frac{1}{\pi a^3} \int dV e^{-(\frac{r_A + r_B}{a})} \equiv S$$

must be taken into account while formulating the eigenvalue problem. Although the overlap integrals can be computed, often analytically, they represent a negative side of this method called "LCAO"-method, LCAO standing for linear combination of atomic orbitals.

Once the Ritz space has been established, the eigenvalue problem of the operator, when cast into the eigenvalue problem for the sought for coefficients C_A and C_B describing the eigenfucntions as linear combination of the two atomic orbitals, reads

$$\begin{pmatrix} H_{AA} - E(R) & H_{AB} - E(R) \cdot S \\ H_{BA} - E(R) \cdot S & H_{BB} - E(R) \end{pmatrix} \begin{pmatrix} C_A \\ C_B \end{pmatrix} = 0$$

The matrix elements write

$$\begin{aligned}
H_{AA} &= <A|H|A>= E_{1s}+e^2/R-\int dV\psi_A^2\frac{e^2}{r_B}\\
H_{BB} &= <B|H|B>= E_{1s}+e^2/R-\int dV\psi_B^2\frac{e^2}{r_A}\\
H_{AB} &= <A|H|B>=-e^2\int dV\psi_A\psi_B\frac{1}{r_A}+E_{1s}\cdot S+\frac{e^2}{R}\cdot S\\
H_{BA} &= <B|H|A>=-e^2\int dV\psi_B\psi_A\frac{1}{r_B}+E_{1s}\cdot S+\frac{e^2}{R}\cdot S
\end{aligned}$$

The various components of the matrix elements are discussed now.

- We observe that $H_{AA}=H_{BB}$ and $H_{AB}=H_{BA}$.
- For building the matrix elements we have taken into account that
$$[-\frac{\hbar^2}{2m}\nabla-\frac{e^2}{r_A}]\psi_A=E_{1s}\psi_A$$
- $-\int dV\psi_A^2\frac{e^2}{r_B}\equiv -V_{AA}$ is the Coulomb interaction of the charge density $-e\psi_A(\mathbf{r})^2$ with the proton B
- e^2/R is the Coulomb repulsion between the protons.
- The integral
$$-e^2\int dV\psi_A\psi_B\frac{1}{r_A}\doteq -V_{AB}$$
is the amplitude that an electron centered at B is transferred to A by the Coulomb interaction originating from the A proton. This is a ”hopping” matrix element which has no counterpart classically.

The eigenvalue problem resulting can be written as ($\xi\equiv E(R)-E_{1s}-\frac{e^2}{R}$)

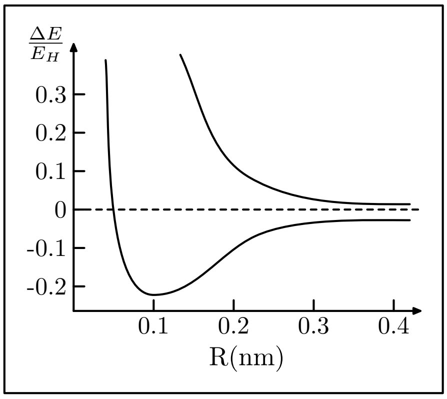
$$(V_{AA}+\xi)C_A+(V_{AB}+\xi\cdot S)C_B=0$$
$$(V_{AB}+\xi\cdot S)C_A+(V_{AA}+\xi)C_B=0$$

Solving the determinantal equation gives the two eigenvalues and eigenfunctions

$$\xi_b(R)=-\frac{V_{AA}+V_{AB}}{1+S}\quad,\quad C_A=C_B=(2(1+S))^{-1/2}$$
$$\xi_a(R)=-\frac{V_{AA}-V_{AB}}{1-S}\quad,\quad C_A=-C_B=(2(1-S))^{-1/2}$$

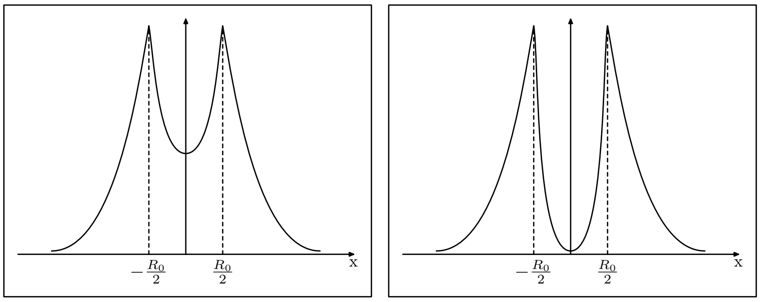
Comments

- The energies associated with ξ_a and ξ_b are shown as a function of R in the figure. Because $V_{AB}>0$ the lower energy of the system belongs to the b-state. In addition, the b-state energy undergoes a minimum at a certain distance R_0. At this distance there is also an energy gain of the molecule against the free atoms: this sanctions the building of the chemical bond.



Computed $\Delta E \equiv E_{b,a} - E_{1s}$ as a function of R ($E_H = -13.6$ eV), redrawn from Feynman lectures on Physics, Vol.3.

- The b-eigenstate ($C_A = C_B$) is called the ”bonding” state, the a-eigenstate is the antibonding one. The possibility of sharing the electron in virtue of the hopping integral renders the chemical bond possible.



Left: the b-and a-wave functions along the axis hosting the two protons. The b-wave function has a strong charge accumulation halfway between the protons. right: The a-wave function has a knot exactly there.

- This chemical bond is essentially due to the non-diagonal matrix element V_{AB}. The hopping process produces the attractive interaction necessary to bind the electron. In the a-state the Coulomb repulsion of the protons dominates.
- This one-electron bond is a (computationally simple) model example for a covalent chemical bond and forms the basis for the independent particle model of chemical bonding: each bonding orbital is able to host two electrons at the most – provided they have opposite spin, in agreement with the Pauli principle. The entire electron configuration of a molecule is obtained by filling successively bonding states, formed by atomic orbitals, with a pair of electrons.
- In the independent model of chemical bonding one can divide the electrons of any atom into two groups: valence electrons (”unpaired ”) occupying the outer shells and paired electrons (core electrons). The number of extra unpaired electrons in one given state of an atom determines its chemical valence. Core electrons are not involved in the formation of the bonding. The outer

shells, instead, provide the bonding states. As there are a finite number of valence electrons, and as their orbitals can be filled with two electrons at the most, a saturation of the chemical forces is achieved. Thus, quantum mechanics and the Pauli principle provide the justification of the usual representation of a molecule as an assembly of atoms, which are linked by a finite number of localized valence strokes – one bond per stroke, each bond containing two electrons.

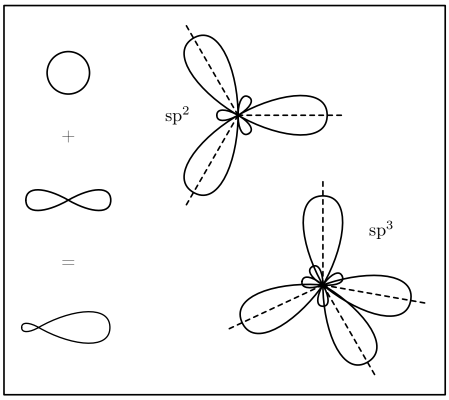
Examples. In the following some examples of chemical bonds are illustrated. In each case the chemical bond is realized as a superposition of atomic-like orbitals that are filled with two electrons so that each atom has complete orbitals.

1. The N_2-molecule. Atomic N has the configuration $(1s)^2(2s)^2(2p)^3$. The wave functions of the three p -electrons of one atom have the possibility of overlapping with the wave function of the p-electrons of the other atom. The process of wave functions overlapping is essential to create a b-orbital and thus for providing the chemical bond. This overlap occurs "head-to-head" along one direction (e.g along z, using the p_z orbitals): the head-to-head bonding is called a σ bond. The p_x-and p_y- orbitaly superposes "side-to-side", i.e. they give rise to two π bondings. The bonding symbol for the $N-2$ molecule is therefore $N \equiv N$.

2. Diatomic carbon C_2. Each C atom has the atomic configuration $(1s)^2(2s)^2(2p)^2$. For forming C_2 only the two p-orbital are needed. One undergoes a σ bond with the other C atom, the second p-electron undergoes a π- bond: $C = C$.

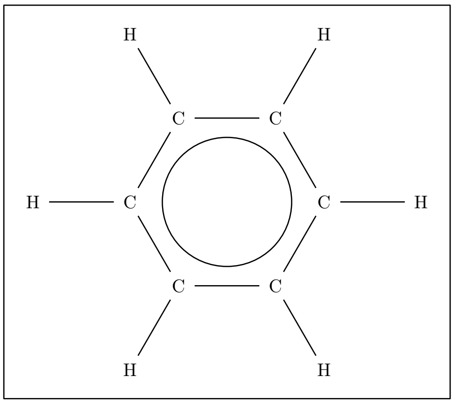
3. Carbon is the element that has the largest chemical diversity. This diversity arises from the fact that 2s and 2p orbitals are very close in energy and can be considered, from the point of view of chemical bond, as almost degenerate. As such, they can be hybridised to form more complex orbitals than p or s orbitals. For instance, the spherical harmonics Y_0^0 and Y_1^0 can be superposed to create an orbital with a particularly directed charge distribution (see Figure). The rationale behind the hybridization is the energy required to build the hybridized orbital can be overcome and more than compensated by the more stable chemical bond achieved by the hybridized orbital. A possible hybridization scheme is e.g. (see Figure)

$$\begin{aligned}
\psi_1 &= \varphi_{pz} \\
\psi_2 &= \frac{1}{\sqrt{3}} \cdot \varphi_s + \frac{\sqrt{2}}{\sqrt{3}} \cdot \varphi_{px} \\
\psi_3 &= \frac{1}{\sqrt{3}} \cdot \varphi_s - \frac{\sqrt{1}}{\sqrt{6}} \cdot \varphi_{px} + \frac{\sqrt{1}}{\sqrt{2}} \cdot \varphi_{py} \\
\psi_4 &= \frac{1}{\sqrt{3}} \cdot \varphi_s - \frac{\sqrt{1}}{\sqrt{6}} \cdot \varphi_{px} - \frac{\sqrt{1}}{\sqrt{2}} \cdot \varphi_{py}
\end{aligned}$$



Left: a s-orbital superposes "head-to-head" with a p-orbital to produce an hybridized orbital. This is an example of a σ bond. Top right: The hybridization of an s-orbital with two p-orbital produces the so-called sp^2 planar hybridized orbitals. Bottom right: The hybridization of an s-orbital with three p-orbital produces the so-called sp^3 tetragonally arranged hybridized orbitals.

This scheme produced three in-plane orbitals forming an angle of 120° and a pz-orbital perpendicular to the plane. This hybridization is used e.g. to build the benzene molecule (see the three strokes originating from the C-atoms representing the σ-bonds arising from the ψ_2, ψ_3, ψ_4 hybridized wave functions. The fourth valence electron in each C atom is in the p_z state perpendicular to the molecular plane. These electrons undergo π-type bonds. The particular symmetry of the molecule produce a complete delocalization of the π electrons, a situation which is described by a ring-like stroke.



The symbolic representation of the benzene molecule.

Because of this delocalization, the six π electrons in the benzene ring can move around almost freely. In fact, when a magnetic field perpendicular to the plane of the benzene ring is switched on an electric circulating current appears in the molecule, causing a magnetic moment. As the current covers a relatively large area, the resulting magnetic moment is large. A large group of other organic compounds have similar molecules like the benzene, the so-called aromatic compounds.

4. The H_2O-molecule. Oxygen has 4-p electrons but only two are actually necessary for binding the two Hydrogen atoms. The molecule show a geometrical configuration with an angle of 105°. If only the two p-electron were involved the angle should be 90°. A larger angle means that the 120° hybridization is also involved in the chemical bond. There is also some transfer charge: the water molecule has an electric dipole.

5. A further way of creating hybridized orbitals is given by the following linear combinations:
$$\begin{aligned}\chi_1 &= \frac{1}{\sqrt{4}}\left(\varphi_s + \varphi_{px} + \varphi_{py} + \varphi_{pz}\right)\\ \chi_1 &= \frac{1}{\sqrt{4}}\left(\varphi_s + \varphi_{px} - \varphi_{py} - \varphi_{pz}\right)\\ \chi_1 &= \frac{1}{\sqrt{4}}\left(\varphi_s - \varphi_{px} + \varphi_{py} - \varphi_{pz}\right)\\ \chi_1 &= \frac{1}{\sqrt{4}}\left(\varphi_s - \varphi_{px} - \varphi_{py} + \varphi_{pz}\right)\end{aligned}$$
These hybridized orbitals point along the four directions of a tetrapod structure and build an angle of 109.6°. This hybridization is e.g. encountered in the methane molecule CH_4. It also builds the building unit of solids with diamond and zinc-blende structure (Si, $GaAs$,...).

9.2. Transition to the solid state

The linear combination of atomic orbitals and the shell model to fill the energy levels can be generalized to a N-atomic molecule and then to a solid ($N \to \infty$). As a way of illustration, we consider the H_N^n molecule on a ring, n being the number of electrons hosted by the molecule. For the computation of the single-electron energy levels we consider the N basis states built of suitable atomic orbitals centered at the various atoms. For simplicity, we assume that the basis states have been orthogonalized so that the overlap integral is set to 0. The ground state energy of the single electron is set to be E_0 – the energy of the electron in one of such basis states. We allow for some hopping amplitude $-A$ so that the eigenvalue problem of the single-electron reads
$$\begin{pmatrix} E_0 - E_i & -A & 0 & \ldots & 0 & -A \\ -A & E_0 - E_i & -A & \ldots & 0 & 0 \\ 0 & -A & E_0 - E_i & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & E_0 - E_i & -A \\ -A & 0 & 0 & \ldots & -A & E_0 - E_i \end{pmatrix} \begin{pmatrix} C_1^i \\ C_2^i \\ C_3^i \\ \ldots \\ C_{(N-1)}^i \\ C_N^i \end{pmatrix} = 0$$
The ring-shaped arrangement allows not only the use of periodic, Born-von Karman boundary conditions, but avoids the occurrence of boundary atoms and with them of unsaturated valencies that would provide an increase in energy. Because of the periodic boundary conditions one can use symmetry arguments to solve the problem: instead of solving with brute force the determinantal equation, we try an Ansatz for the eigenfunctions themselves. The translational symmetry of the problem suggest that eigenfunctions have this property:
$$| C_1^i |^2 = ... = | C_N^i |^2 = \frac{1}{N}$$
This property is fulfilled by setting e.g.
$$C_j^i = \frac{1}{\sqrt{N}} e^{ik^i \cdot a \cdot j}$$

a being the atomic distance and k^i a real number labelling the eigenstate ψ^i. Because $C_{j+N} = C_j$

$$e^{ik^i a(j+N)} = e^{ik^i aj}$$

which determines

$$k^i = \frac{2\pi}{a \cdot N} \cdot i (i = 1, ..., N)$$

k^i is a quantum number that labels the eigenvalues. The eigenfunction belonging to k_i writes

$$\psi^i = \frac{1}{\sqrt{N}} \sum_j e^{ik^i ja} \psi_j$$

This expression generalizes the linear combination of atomic orbitals that we have already encountered in the H_2^+ molecule. The eigenvalues to k_i can be obtained from

$$\frac{1}{\sqrt{N}} [E_0 - E_i] \cdot e^{ik^i a} - \frac{1}{\sqrt{N}} \cdot A \cdot e^{ik^i 2a} - \frac{1}{\sqrt{N}} \cdot A \cdot e^{ik^i aN} = 0$$

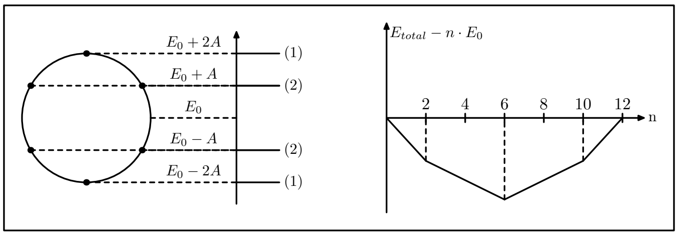
with solution

$$E_i = E_0 - 2 \cdot A \cdot \cos(k^i \cdot a)$$

($k^i = \frac{2\pi}{N \cdot a} \cdot i$). The atomic level E_0 spreads out to a set of discrete molecular levels labelled by k_i. E_0 is the center of mass of the molecular levels. These are the single-particle molecular eigenvalues that provide the eigenstates to be filled with an appropriate number of electrons. The shell model requires that each level can be filled with two electrons. When the lower lying levels are complete, the next ones are filled. Each atomic level $E_1 > E_0, E_2 > E_1, ...$ is bound to provide a set of such discrete molecular levels with spread around the corresponding atomic level.

Examples

- In the figure we plot the total binding energy $E_i - E_0$ as a function of the filling number n for the H_6-molecule.

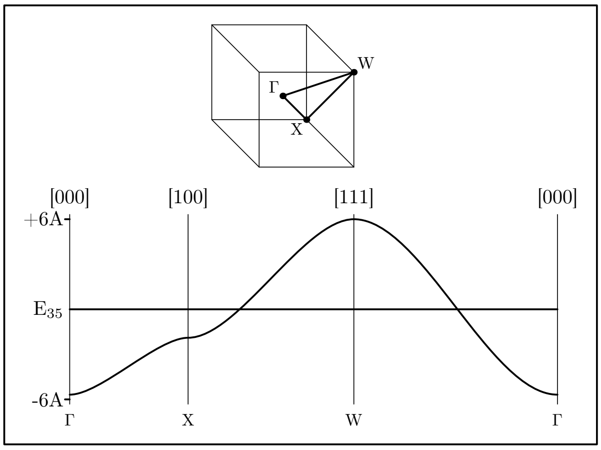


Left: Term scheme of one electron in a H_6 proton ring. Right: total binding energy as a function of the filling number n.

From the diagram left we see that two of the six eigenvalues are twice degenerate, i.e. have the same energy and different k^i. Two are simply degenerate (including the lowest energy level). Let us construct the shells for $n = 6$. We start by placing two electrons at the lowest level, $E_0 - 2A$. Thereafter, four electrons can be accommodated in the level $E_0 - A$, reaching the $n = 6$ filling factor. Higher lying levels can be used to obtain excited states or to continue filling with electrons. The total energy of the ground state of the molecule is $6E_0 - 8A$. Compared with the total ground state energy of the separated atoms ($6E_0$), the chemical bond has produced a gain in energy of $-8A$, which we call the binding energy. The right-hand side of the figure plots the binding energy as a function of n. The neutral H_6 molecule is the most stable form of the molecule though H_6^{4+} and H_6^{4-} are also more stable than others. The reason for this is simple:

with $n = 2, 6, 10$ shells are completely filled, and that's always a reason for enhanced stability, as we know from the noble gas configuration. A fill factor of 12 electrons does not lead to a stable chemical bond, in this approximation.

- Even for nuclei one can construct a shell model, although the nuclear forces are very strong. Experimentally, one has, indeed, found particularly stable nuclei, with characteristic, particularly stable number of nucleons, given by the magical numbers $2, 8, 20, 28, 50, 82$. The idea that the electronic structure of nuclei could be explained by a simple shell model with appearance of "magic numbers" originated with Maria Mayer, which received the nobel price for it.
- The transition to a one-dimensional solid is obtained by generalizing the expression
$$E_i^j = E^j - 2 \cdot A_j \cdot \cos(k^i \cdot a)$$
($k^i = \frac{2\pi}{N \cdot a} \cdot i$) to large N. E^j is the atomic level that spreads out to a band of energy levels. The k-values are very close to each other and occupy a segment $[-\frac{\pi}{a}, \frac{\pi}{a}]$ along the k-axis. The segment is known as the first Brillouin zone of the one-dimensional lattice and the k-axis is the reciprocal space of the one-dimensional lattice. In virtue of the closeness of the k-values the energy bands are almost continuous. The filling of the energy levels with electrons proceeds according to the Pauli principle. In three dimensions, the first Brillouin zone has a polyhedral shape and the **k** vectors build a dense set within this polyhedron. Energy bands are plotted usually along some direction and at some points within the polyhedron.



Tight-binding s-bands in a simple-cubic structure computed including only nearest-neighbour overlaps. The horizontal line is at the center of gravity of the band. The inset gives the first Brillouin zone of the simple cubic lattice and the points and lines along which the bands are plotted.

9.3. The H_2 molecule

The computation of energy levels for the H_2^+ molecule is the first essential step of the filling of energy states within the independent particle shell model. When the states single-electron states are filled according to the Pauli principle one produces a many-electron system that is physically plausible. However, an accurate quantitative verification of the independent particle model actually requires an explicit two-particle computation that confirms the strategy of filling of states with the Pauli principle. This quantitative proof was the achievement of Heitler and London, who provided a detailed computation of the H_2-molecule that provided an a posteriori justification of the independent particle model for molecules. The Heitler-London computation was based again on the Born-Oppenheimer approximation, where the two nuclei A,B are kept fixed at a distance R which enters the problem as a parameter. However, they used as basis states for the Ritz space the spin-triplet and spin-singlet two-particle states

$$\chi_t(1,2)\otimes\underbrace{[2(1+S^2)]^{-1/2}(\psi_A(1)\otimes\psi_B(2)+\psi_A(2)\otimes\psi_B(1)}_{\varphi_t(R,1,2)}$$
$$\chi_s(1,2)\otimes\underbrace{[2(1-S^2)]^{-1/2}(\psi_A(1)\otimes\psi_B(2)-\psi_A(2)\otimes\psi_B(1))}_{\varphi_s(R,1,2)}$$

with

$$\begin{aligned}\psi_A(1)&=(\pi a^3)^{-1/2}e^{(\frac{-r_{A1}}{a})}\\ \psi_A(2)&=(\pi a^3)^{-1/2}e^{(\frac{-r_{A2}}{a})}\\ \psi_B(1)&=(\pi a^3)^{-1/2}e^{(\frac{-r_{B1}}{a})}\\ \psi_B(2)&=(\pi a^3)^{-1/2}e^{(\frac{-r_{B2}}{a})}\end{aligned}$$

being derived from the basis s-states of the separated atoms. Again, the use of atomic orbitals introduces an (annoying) overlap integral

$$S=\int dV_1\psi_A(1)\psi_B(1)=\frac{1}{\pi a^3}\int dV_1 e^{-\frac{r_{A1}+r_{B1}}{a}}$$

that must be computed to obtain accurate results. The Hamilton operator – neglecting both motion of the nuclei and spin-orbit coupling – writes

$$\hat{H}=\frac{-\hbar^2}{2m}(\triangle_1+\triangle_2)-e^2[\frac{1}{r_{A1}}+\frac{1}{r_{A2}}+\frac{1}{r_{B1}}+\frac{1}{r_{B2}}-\frac{1}{r_{12}}-\frac{1}{R}]$$

The indices 1,2 refer to the electrons, the indices A,B to the nuclei. If one consider that the spin-singlet and spin-triplet states are orthogonal, $\varphi_s\otimes\chi_s$ is orthogonal to $\varphi_t\otimes\chi_t$, within the Ritz space the matrix of $\hat{H}$ is a four-by-four diagonal matrix:

$$\begin{pmatrix}(\varphi_t,\hat{H}\varphi_t) & 0\\ 0 & (\varphi_s,\hat{H}\varphi_s)\end{pmatrix}$$

and there is one threefold degenerate eigenvalue $E_t=(\varphi_t,\hat{H}\varphi_t)$ and one non-degenerate eigenvalue $E_s=(\varphi_s,\hat{H}\varphi_s)$. For calculating the matrix elements one must consider that

$$[\frac{-\hbar^2}{2m}\triangle_1-\frac{e^2}{r_{1A}}]\psi_A(1)=E_{1s}\psi_A(1)$$

The computation of the matrix elements gives the two eigenvalues

$$\begin{aligned}E_t&=2E_{1s}+\frac{Q-J}{1-S^2}\\ E_s&=2E_{1s}+\frac{Q+J}{1+S^2}\end{aligned}$$

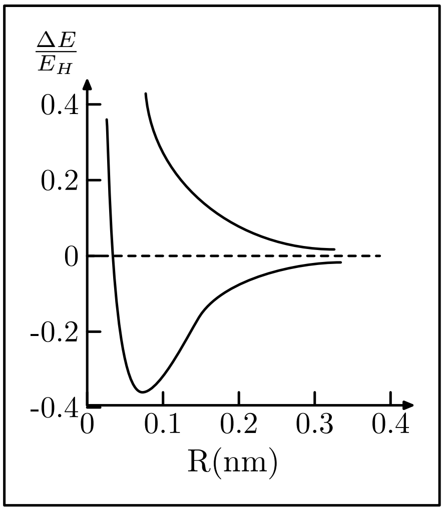
Let us now discuss now the parameters entering the eigenvalues:

$$\begin{aligned}Q&=\int dV_1dV_2\psi_A^2(1)\psi_B^2(2)[\frac{-e^2}{r_{B1}}+\frac{-e^2}{r_{A2}}+\frac{e^2}{r_{12}}+\frac{e^2}{R}]\\ &=-\int dV_1\psi_A^2(1)\frac{e^2}{r_{B1}}-\int dV_2\psi_B^2(2)\frac{e^2}{r_{A2}}+\int dV_1dV_2\psi_A^2(1)\psi_B^2(2)\frac{e^2}{r_{12}}+\frac{e^2}{R}\end{aligned}$$

contains various Coulomb energies, such as the Coulomb energy of the first electron with the nucleus B, the Coulomb energy of the second electron with nucleus A, the Coulomb interaction of the two electrons and the Coulomb repulsion of the two nuclei. Q is called **Coulomb-integral**. The integral

$$J=\frac{e^2\cdot S^2}{R}+\int dV_1 dV_2 \psi_A(1)\psi_B(2)\frac{e^2}{r_{12}}\psi_A(2)\psi_B(1)$$
$$-S\int dV_1\psi_A(1)\frac{e^2}{r_{B1}}\psi_B(1)-S\int dV_2 d\psi_B(2)\frac{e^2}{r_{A2}}\psi_A(2)$$

represents the **exchange energy**, arising from the requirement of anti-symmetrizing the wave functions. Both Q and J depends on the distance R. Both can be computed exactly within the subspace chosen in the H-L method. $Q(R)$ is a small positive quantity. J is instead negative. Accordingly, the singlet state has the lowest energy. Notice that the a negative sign of J is the essential ingredient for a stable chemical bond, which then occurs in the singlet state. However, a negative J leads to a preferred antiparallel coupling of the two spins and is a simple but very clear and robust demonstration that anti-parallel spin alignment is the key coupling mechanism when atoms are assembled to form a solid. We thus have a very complicated situation where the Pauli principle, essential for the chemical bond, works against parallel spin alignment – and ultimately, e.g., against e.g. a ferromagnetic state.



Plot of the $\Delta E=E_s-E_t$ as a function of R. The singlet state has a minimum at a well defined distance R_0, at which the chemical bond of the two H-atoms is realized. R_0 is the equilibrium distance of the nuclei in the the H_2-molecule. The experimental value is $0.7395A°$.

In summary, the two-body computation finds that the two electrons occupying a chemical bond are in a singlet state. When we compare the energies obtained from the independent particle model and from the explicit two-body computation, we find a similarity of the R dependence between 1. on one side the one-particle bonding state and the two-particle singlet state and 2. on the other the one-particle anti-bonding state and the two-particle triplet state. This similarity is explained when we compute the charge density distribution produced by the two-particles wave functions. In the presence of only one electron with wave function $\psi(\mathbf{r})$, the charge density is given by $|\psi(\mathbf{r})|^2$. When the system has two or more particles, one must find a physical way of defining a charge density. This was done by W. Kohn with his "Density functional approach": given the normalized wave function

$\psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$, the charge density $\rho(\mathbf{r})$ is given by

$$\rho(\mathbf{r}) = N \cdot \int dV_2 \cdot dV_3 ... \cdot dV_N \cdot |\, \psi(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r}_N) \,|^2$$

The charge distribution computed with the two-particles orbital singlet wave function also gives the same accumulation of charge halfway between the two protons as observed in the one-electron bonding state – we conclude that it is the accumulation of charge between the two nuclei – and, ultimately, the hopping matrix element – that actually drives the chemical bond. The two-body computation provides the important information that the two electrons entering the chemical bond build a singlet, in agreement with Pauli principle.

Part IV.

Some advanced topics

10. Quantum scattering.

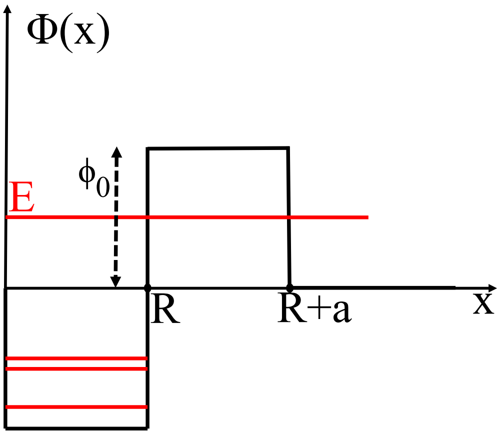
10.1. Introduction

When solving the Schrödinger equation in a spherically symmetric potential, one encounters non-square integrable wave functions, appearing in the continuous range of the spectrum of an Hamilton operator. We will see that they are used to describe the scattering of particles at some potential energy. Such scattering states are encountered e.g. in the description of the quantum mechanical tunnel effect and the diffraction of material waves at crystals: we will treat both situations in the following sections. We point out that the computational technology described in this chapter for purely quantum mechanical system can be translated to the description of e.g the diffraction of classical scalar waves and light, which also fulfill the Helmholtz equation. We emphasize again one particular characteristic of scattering states, that we have seen in the solution of the SE in spherically symmetric potentials: in the range where E assumes continuous values, the SE is not intended as an eigenvalue equation for some sought for eigenvalues E: E is not restricted to any particular discrete value and the SE has solutions for any value of the parameter E.

10.2. G. Gamov (1928): the quantum mechanical tunnel effect

Natural radioactivity, discovered by Becquerel in 1896 and studied in depth by Pierre and Marie Curie, consists of unstable atomic nuclei changing spontaneously into other nuclei (e.g. $U_{92}^{238} \rightarrow Th_{90}^{234} + \alpha_2^4$). In the process, particles and very short-wave radiation is emitted. The emitted particles are e.g. α particles, i.e. He-cores He^{++}. Later, Rutherford discovered the artificial nuclear transformation i.e. the ability, by shelling of particles, to transform nuclei of an element into nuclei of another element. Understanding the α decay of the heavy nuclei represented a challenge that is beyond the reach of classical physics and was solved by quantum mechanics. The description of this decay provides a model for any process that involves quantum mechanical tunneling.

For the sake of illustration, we consider the quantum mechanical motion of the charged He^{++} particle with mass m to be along the x-axis, subject to the a potential energy that entails bound states between 0 and R (potential well) and scattering state between $R+a$ and ∞. Between R and $R+a$ there is a rectangular potential barrier.

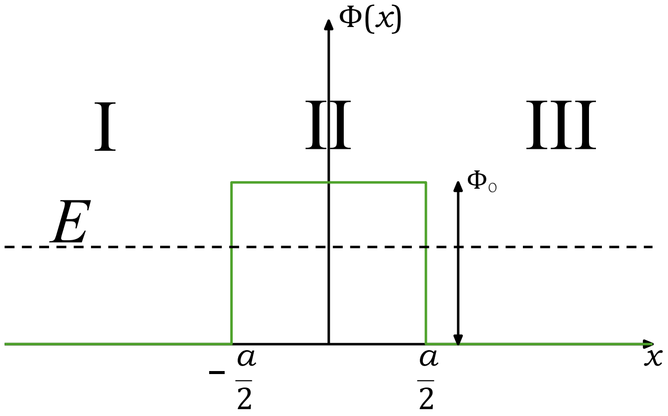


Schematic potential barrier in Gamov tunneling. The α *particle in in a bound state of the nucleous, its energy level is given as E (red line). The potential barrier is* $\Phi_0 > E$*. Accordingly, the particle cannot penetrate, in classical mechanics, the range* $[R, R+a]$.

The α particle is, initially, caged within the nucleus ($x \in [0,R]$) at some energy level E. At R the potential energy assumes the role of a ”wall’. According to classical mechanics, a particle impinging on the wall from the left is reflected at the potential barrier if the energy is less than Φ_0. For $E > \Phi_0$ the particle is transmitted to the positive x-axis. In other words, the potential wall is completely transparent for the particle for $E > \Phi_0$ and a perfect mirror for $E < \Phi_0$. A classical particle is EXCLUDED from the interval $x \in [R, R+a]$ because there the kinetic energy $E_{kin} = E - \Phi_0$ would be negative. In quantum mechanics, G. Gamov found a different behaviour.
For obtaining Gamov solutions, we solve explicitly the SE

$$\frac{d^2}{dx^2}\psi(x) + \frac{2m}{\hbar^2}\cdot(E - \Phi(x))\cdot\psi(x) = 0$$

for $E < \Phi_0$ in a situation where the potential has a square shape.



Rectangular potential barrier with width a and height Φ_0. We are interested at the solutions of the Schrödinger equation for a state with energy $E < \Phi_0$.

This is a homogeneous ordinary differential equation, but one where the coefficients depends on the independent variable x through the potential energy $\Phi(x)$. We have already encountered this mathematical situation in the case of the quantum mechanical harmonic oscillator and in the infinite potential well problem. We have learned by solving the problem of the ”particle in box” that quantum mechanical problems with this type of potential can be solved by finding solutions of the SE within a certain interval (where it can be solved exactly) and then requiring these solutions (and their derivative) to be continuous at the points that divide the intervals. This continuities are necessary for the SE to be mathematically robust. For a potential wall going to infinity, e.g., the boundary condition requires that the wave function vanishes at the point where the potential sets in. There, it goes to infinity and the term $\Phi(x)\cdot\psi(x)$ occurring in the Schrödinger equation is only finite if $\psi(x)$ is vanishing. A rectangular potential barrier has two points where the potential is non-continuous. Mathematically speaking, this discontinuity is translated to the Schrödinger equation, again through the term $\Phi(x)\cdot\psi(x)$. It must be compensated by a discontinuity in the second derivative of the wave function. A discontinuity in the second derivative implies that the wave function itself and its first derivative are, at least, continuous. Accordingly, for a rectangular potential barrier we require the continuity of the wave functions and of its derivatives at the site where the barrier makes a jump.
We find first the solutions of the SE in the ranges (I, II, III):

$$\begin{aligned}\varphi_I(x) &= Ae^{i\alpha x} + Be^{-i\alpha x}\\ \varphi_{II}(x) &= Ce^{\beta x} + De^{-\beta x}\\ \varphi_{III}(x) &= Fe^{i\alpha x} + Ge^{-i\alpha x}\end{aligned}$$

mit $\alpha \equiv \sqrt{2mE/\hbar^2}$ und $\beta \equiv \sqrt{2m(V_0 - E)/\hbar^2}$. In the range $E < V_0$, the wave functions are not normalizable, so that the total wave functions is not normalizable. A proper way to proceed would be to introduce walls at distant points and install periodic boundary conditions. Here we adopt a different way: we dispense with finding a normalization constant and try to find results where an hypothetical normalization constant would cancel out. The boundary conditions produce a system of coupled equations for the sought-for constants A to F:

$$\begin{aligned}
Ae^{-i\alpha a/2}+Be^{i\alpha a/2} &= Ce^{-\beta a/2}+De^{\beta a/2} \\
i\alpha(Ae^{-i\alpha a/2}-Be^{i\alpha a/2}) &= \beta(Ce^{-\beta a/2}-De^{\beta a/2}) \\
Ce^{\beta a/2}+De^{-\beta a/2} &= Fe^{i\alpha a/2}+Ge^{-i\alpha a/2} \\
\beta(Ce^{\beta a/2}-De^{-\beta a/2}) &= i\alpha(Fe^{i\alpha a/2}-Ge^{-i\alpha a/2})
\end{aligned}$$

As we investigate particles that originate on the left hand side and propagate to the right, we can set the coefficient of the plane wave travelling from right to left to zero, i.e. $G=0$. We obtain a system of four equations for the 5 coefficients A, B, C, D, F, from which we can compute the ratios B/A, C/A, D/A, and F/A ($A \neq 0$). Notice that, by building these ratios, we eliminate hypothetical normalization constants from the problem. In the following, we will need only $|B/A|^2$ and $|F/A|^2$, which we give here explicitly:

$$\begin{aligned}
\frac{|B|^2}{|A|^2} &= \frac{1/4(\alpha/\beta+\beta/\alpha)^2\sinh^2(2\beta\frac{a}{2})}{1+1/4(\alpha/\beta+\beta/\alpha)^2\sinh^2(2\beta\frac{a}{2})} \\
\frac{|F|^2}{|A|^2} &= \frac{1}{1+1/4(\alpha/\beta+\beta/\alpha)^2\sinh^2(2\beta\frac{a}{2})}
\end{aligned}$$

In Schrödinger quantum mechanics (see the next chapter for a full explanation) one can compute, starting from the wave function, a current density by means of the expression

$$\vec{J}(\vec{r},t) \doteq \frac{\hbar}{2im}\Big(\overline{\psi(\vec{r},t)}\vec{\nabla}\psi(\vec{r},t)-(\vec{\nabla}\overline{\psi(\vec{r},t)})\psi(\vec{r},t)\Big)$$

The currents in the regions I, II, III amount to

$$\begin{aligned}
J_I &= \frac{\hbar\alpha}{m}(\bar{A}A-\bar{B}B)=J_I^e-J_I^r \\
J_{II} &= \frac{\hbar\beta}{im}(\bar{D}C-D\bar{C}) \\
J_{III} &= \frac{\hbar\alpha}{m}\bar{F}F
\end{aligned}$$

where J_I^e is the current density of a wave hat runs from left to right, J_I^r is the current density of a wave that is reflected at the wall and J_{III} is the current density of the wave that is transmitted into region III. Using J_I^e, J_I^r und J_{III} we compute two measurable coefficients which are independent of the putative normalization constants: the reflection coefficient

$$R \equiv \frac{J_I^r}{J_I^e} = \frac{|F|^2}{|A|^2} = \frac{1/4(\alpha/\beta+\beta/\alpha)^2\sinh^2(2\beta\frac{a}{2})}{1+1/4(\alpha/\beta+\beta/\alpha)^2\sinh^2(2\beta\frac{a}{2})}$$

and the transmission coefficient

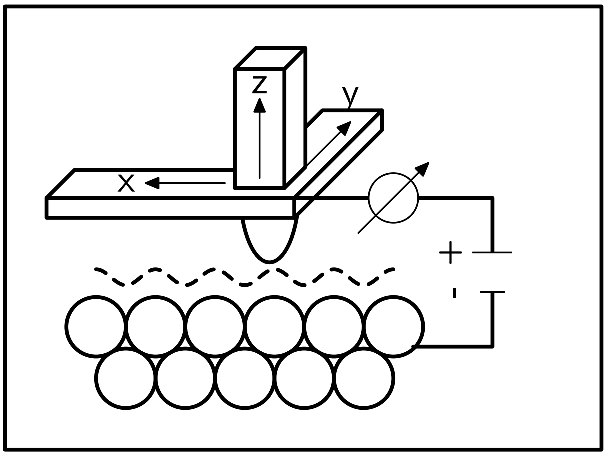
$$D \doteq \frac{J_{III}}{J_I^e} = \frac{|F|^2}{|A|^2} = \frac{1}{1+1/4(\alpha/\beta+\beta/\alpha)^2\sinh^2(2\beta\frac{a}{2})}$$

with $R+D=1$. $D \neq 0$ means that a particle that arrives from the left has a finite probability of passing through a classically forbidden zone II and being transmitted into the region III beyond the potential wall. One speaks of ”tunnel effect”. For large a we obtain an eponential decay of the tunneling probability:

$$D \propto e^{-4\cdot\pi\frac{a}{\Lambda}}$$

where $\lambda = \sqrt{\frac{\hbar^2}{2m(\Phi_0-E)}}$ is the De Broglie wave length for a particle with energy Φ_0-E. Some limiting cases are interesting: if the mass is very large (macroscopic objects) or $\hbar \rightarrow 0$ (classical limit), $D \rightarrow 0$, i.e. the particle cannot enter the classically forbidden region inside the potential barrier and is totally reflected. For an electron ($m=0.9\cdot 10^{-30}kg$) with 1 eV energy, a barrier height of 4 eV produces a characteristic length of about 0.6 nm. At this length, the transmission coefficient is reduced by a factor of e^{-1}. Sizeable electron currents are, accordingly, still obtained in artificial lattices with barrier width in the nm range. Protons and heavier particles have much larger mass and therefore a much smaller transmission coefficient.

This simple model calculation of the tunnel effect makes numerous physical phenomena at atomic level understandable in principle. These include very important ones like the escape of nucleons and nucleon complexes from an atomic nucleus in the radioactive decay, the cold electron emission of metals in strong electric fields (field emission) or the passage of electrons through oxide layers between metal contacts (diode and transistors), and the operating of semiconductor multilayers. A modern application of the tunnel effect as a method of studying the morphology of metallic surfaces with atomic resolution is the **Scanning Tunnelling Microscope**.



With the help of a piezoelectric actuator a tip made of platinum/iridium or tungsten is moved along z and brought at about 1 nm distance from the surface of an electrically conductive sample. At this distance, the electron wave function of the surface atoms overlap with those of the tip. Because of the small distance, the electrons can tunnel between tip and sample atom through the vacuum barrier. This produces a tunnel current, the strength of which depends on the distance between tip and sample (the parameter a in our computation) and thus, by reading the tunnel current one can obtain a height relief of the scanned surface. By moving the tip parallel to the sample surface (x — y-plane) and recording the tunnel current an image of the sample surface can be generated, which clearly shows the atomic structure of the surface.

10.3. *The general treatment of scattering states

We write the SE as an inhomogeneous Helmholtz equation

$$\triangle_q \psi(q) + \underbrace{\frac{2m \cdot E}{\hbar^2}}_{k^2} \psi(q) = \frac{2m}{\hbar^2} \cdot \Phi(q) \cdot \psi(q)$$

and solve it with an algorithm which is used to solve quite general linear inhomogeneous differential equations occurring in both classical and/or quantum physics.

10.3.1. The Green function method

The Green function The Green function method, widely used in physics and engineering, is an algorithm designed to find one particular solution of equations of the type

$$\mathscr{L}u(p) = f(p)$$

where $\mathscr{L}$ is some linear differential operator, and $f(p)$ is a function that represents an inhomogeneous localized "source" of perturbation. The symbol "p" represents a set of variables like the cartesian coordinates (x, y, z). The problem is defined in a region of p-space that we call D and the sought-for solution must fulfill some boundary condition along the surface δD. What we have in mind is e.g. the Poisson problem

$$\triangle_p u(p) = \rho(p)$$

where $\rho(p)$ is a localized charge distribution and the sougth for solution $u(p)$ is the potential of the charge distribution. The SE written in the form quoted above is also in this class of problems. The Green function method consists in replacing the source $f(p)$ with the point source $\delta(p-q)$, based at the soutce point q. The Green function $G(p,q)$ of the problem is that linear superposition of a particular solution of the inhomogeneous problem

$$\mathscr{L}u(p,q) = \delta(p-q)$$

with the general solution of the homogeneous problem which fulfills the boundary conditions.
We now show that, knowing the Green function of the problem, generates the sought for solution of the original problem, defined for a more extended source $f(p)$. We Start from

$$\mathscr{L}G(p,q) = \delta(p-q)$$

Multiply both sides with $f(q)$

$$\mathscr{L}G(p,q)\cdot f(q) = \delta(p-q)\cdot f(q)$$

Integrate both sides over the domain Ω defined by the localized source

$$\int_\Omega d\Omega_q \mathscr{L}G(p,q)\cdot f(q) = \int_\Omega d\Omega_q \delta(p-q)\cdot f(q)$$

By virtue of the properties of the δ-function, the right-hand side gives just $f(p)$. As $\mathscr{L}$ is a linear operator and depends only on the variable p, we can transfer it before the integral and obtain

$$\mathscr{L}\int_\Omega d\Omega_q G(p,q)\cdot f(q) = f(p)$$

from which we can read out that

$$\int_\Omega d\Omega_q G(p,q)\cdot f(q)$$

is the sought for solution of the original inhomogeneous problem.

EXAMPLES.
1. As an application, we find the Green function of the Poisson problem

$$\triangle\Phi(p) = -\frac{\rho(p)}{\epsilon_0}$$

in the three-dimensional space, i.e. the solution of

$$\triangle G(p,q) = \delta(p-q)$$

p, q being some vectors $\mathbf{r}_p, \mathbf{r}_q$ and $p-q = \mathbf{r}_p - \mathbf{r}_q$. As a boundary condition we require that the differential equation is defined on a domain $D = \mathscr{R}^3$ with the vanishing of the solution at $\mid p-q \mid \to \infty$. For constructing an Ansatz for $G(p,q)$ we notice that for $p \neq q$ the differential equation is the Laplace equation $\triangle G(p,q) = 0$ which is solved $e.g$ by $\frac{1}{|p-q|}$. The Ansatz for the Green function is therefore

$$G(p,Q) = c\cdot\frac{1}{\mid p-q \mid}$$

c being a to be determined constant. We determine c by requiring that

$$\int_D dV_p \cdot \triangle_p (c \cdot \frac{1}{|p-q|}) t(p) = t(q)$$

for any test-function defined over the domain D. To compute the integral on the left-hand side using Green formula to transfer the Laplace operator onto the test-function:

$$\int_D dV_p (\triangle G \cdot t(p)) = \int_D dV_p G(p,q) \triangle t(p) + \underbrace{\int_{\partial D} dO_p (\frac{\partial G}{\partial n} \cdot t(p) - \frac{\partial t}{\partial n} \cdot G(p,q)}_{\to 0}$$

As the surface ∂D is pushed to infinity, where the functions and their normal derivative vanishes, the surface integrals vanish. We now set q at the origin of the coordinate system, take spherically symmetric test-functions and

$$\triangle t(p) = \frac{1}{r^2} \frac{d}{dr} (r^2 \frac{dt(r)}{dr})$$

Inserting $G(r) = \frac{c}{r}$ and performing the integration over the angular variable gives

$$\int_D dV_p G(p,q) \triangle t(p) = 4\pi \cdot c \cdot \int_0^\infty r^2 dr \frac{1}{r} \cdot \frac{1}{r^2} \frac{d}{dr} (r^2 \frac{dt(r)}{dr}$$

Partial integration gives

$$\int_D dV_p G(p,q) \triangle t(p) = 4\pi \cdot c \cdot \int_0^\infty \frac{dt}{dr} dr = -4\pi \cdot c \cdot t(q)$$

Requiring that this integral being equal $t(q)$ leads to

$$c = -\frac{1}{4\pi}$$

Using this result we can write down the potential $\Phi(p)$ emanating from a localized charge distribution $\rho(p)$:

$$\Phi(p) = \frac{1}{4\pi\epsilon_0} \int_{D(\rho(q))} dV_q \frac{1}{|p-q|} \cdot \rho(q)$$

This is the superposition principle of electrostatics.

2. We now turn to the inhomogeneous Helmholtz equation

$$\triangle G(p,q) + k^2 \psi(p) = \delta(p-q)$$

that we want to solve in $D = \mathscr{R}^3$ with the boundary condition that the wave is propagating away from the source in q but without specifying any other boundary conditions. The Ansatz for finding the Green function is again obtained by using the one solution of the homogeneous equation that is fullfilling the inhomogeneous Helmoltz equation with point like source "almost" everywhere:

$$G(p,q) = c \cdot \frac{e^{i \cdot k \cdot (|p-q|)}}{|p-q|}$$

The constant c is then determined by evaluating the test-integral and turns out to be again $-\frac{1}{4\pi}$. This is not surprising as the Green function of the Helmholtz equation is the Green function of the Poisson equation for $k = 0$. The Green function is then the superposition of this particular solution with the general solution of the homogeneous Helmholtz equation. By superposing the general solution of the homogeneous equation one takes care of fulfilling the boundary conditions.

10.3.2. Application to potential scattering

We have now the tools to find a scattering solution of the SE

$$\triangle_{\mathbf{x}} \psi(\mathbf{x}) + k^2 \cdot \psi(\mathbf{x}) = \frac{2m}{\hbar^2} \cdot \Phi(\mathbf{x}) \cdot \psi(\mathbf{x})$$

$\Phi(\mathbf{x})$ is some potential that we consider as ”localized”. Localized means that it has a domain Ω into which it has sizeable values. Close to the boundary of this domain it decays rapidly to zero. This is not the case of e.g. the Coulomb potential, but take the attitude that as long as the various integrals appearing in the solution converge one can use the solution itself. We are seeking a particular Green function of this problem which can be used to describe scattering states, so that we impose the boundary condition that the Green function decays for $|\mathbf{x}| \to \infty$ – which is a physically plausible solution for a localized potential. In addition, we seek a Green function that describe a wave travelling away from the source. The sought for Green function reads

$$G(\mathbf{x},\mathbf{y}) = -\frac{1}{4\pi} \cdot \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}$$

The scattered wave originating from the ”source” $\frac{2m}{\hbar^2} \cdot \Phi(\mathbf{x}) \cdot \psi(\mathbf{x})$ and decaying to 0 sufficiently away from Ω writes

$$\psi_s(\mathbf{x}) = -\frac{2m}{4\pi\hbar^2} \int_\Omega d^3y \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|} \cdot \Phi(\mathbf{y}) \cdot \psi(\mathbf{y})$$

The integral on the right-hand side is reminescent of the Huygens principle (1660) for the finding of the diffracted wave at an aperture: ”The wave field behind an aperture is made up of the coherent superposition of spherical waves, the origin of the spherical waves being the individual geometric points in the aperture itself”. The general solution of the SE in the presence of the potential is the superposition of the scattered wave with a solution of the homogeneous SE that describes the proper asymptotic behaviour at large distances from Ω. There, we require the wave to be a plane wave, as we want to describe the potential scattering of plane waves, so that the solution of the SE for plane wave scattering writes

$$\psi(\mathbf{x}) = e^{ik\cdot\mathbf{n}_0\cdot\mathbf{x}} - \frac{2m}{4\pi\hbar^2} \int_\Omega d^3y \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|} \cdot \Phi(\mathbf{y}) \cdot \psi(\mathbf{y})$$

Looking more closely to this ”solution”, we find that the sought for wave appears both on the left-hand side and within the integral. So, this is not a solution but an integral equation for the sought for wave – it is called the Lippman Schwinger equation. But put in this form, we are able to propose a simple approximation for the computing of the integral on the right-hand side that produces indeed a ”solution”, albeit an approximate one. We are guided by this from the approximation introduced by Kirchoff and Fresnel to solving the ”Huygens” integral for diffraction: instead of using the exact solution $\psi(\mathbf{y})$ in the region Ω on the right-hand side, we plug in the plane wave itself. The plane wave is only an approximation for the true wave in the region Ω, but its use makes the integral computable by elementary means. In quantum mechanics, this approximation is called the Born approximation of the scattering solution. The first step is just giving the approximate total wave as

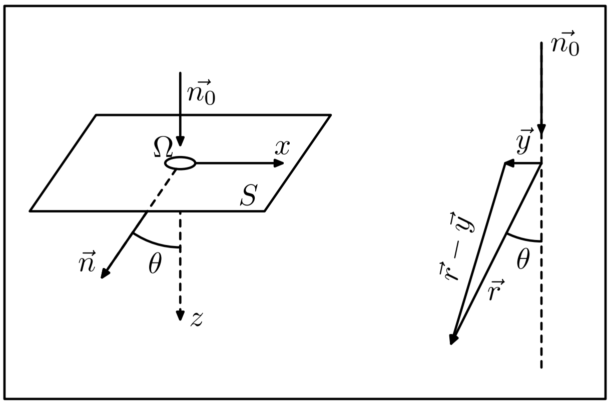
$$\psi(\mathbf{x}) \approx e^{ik\cdot\mathbf{n}_0\cdot\mathbf{x}} - \frac{2m}{4\pi\hbar^2} \int_\Omega d^3y \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|} \cdot \Phi(\mathbf{y}) \cdot e^{ik\cdot\mathbf{n}_0\cdot\mathbf{x}}$$

The second step is that we need to know the wave far away from the source – detectors are usually placed far away from the scattering target. In the far region

$$|\mathbf{r}-\mathbf{y}| \approx |\mathbf{r}| - \mathbf{y}\underbrace{\frac{\mathbf{r}}{|\mathbf{r}|}}_{\mathbf{n}}$$

and the approximate wave there writes

$$\psi(\mathbf{x}) \approx e^{ik\cdot\mathbf{n}_0\cdot\mathbf{x}} + \underbrace{\frac{e^{ik|\mathbf{r}|}}{|\mathbf{r}|} \cdot \left(-\frac{2m}{4\pi\hbar^2} \int_\Omega d^3y\, e^{ik\cdot(\mathbf{n}_0-\mathbf{n})\cdot\mathbf{y}} \cdot \Phi(\mathbf{y})\right)}_{\psi_s(\mathbf{x})}$$



Left: a typical situation encountered in scattering. The plane wave is impinging from above along the $+z$*-direction onto the plane* S *where the region of localized potential* Ω *is located. The scattered intensity is observed e.g. on the* zx*-plane for a varying angle* θ *with respect to the forward direction* $+z$*. Right: Cross section of the region* Ω *along the* zx*-plane. A spatial location within the region* Ω *is identified by the vector* $\mathbf{y}$*. The scattered wave is computed at a site indicated by* $\mathbf{r}$*.*

The amplitude of the outgoing spherical wave is the so-called scattering amplitude $f(k,\mathbf{n}_0,\mathbf{n})$ and it is, in Born approximation

$$f(k\cdot(\mathbf{n}_0-\mathbf{n}))_{\text{Born}} = -\frac{2m}{4\pi\hbar^2}\int_\Omega d^3y e^{ik\cdot(\mathbf{n}_0-\mathbf{n})\cdot\mathbf{y}}\cdot\Phi(\mathbf{y})$$

This result is particularly simple to remember, because it associates the scattering amplitude with the Fourier transform of the potential.

Having constructed the wave in the far region we ask ourselves what will be the result of an attempt at detecting the scattered particles. We imagine a detector placed at large distances from the scattering target and oriented in such a way that it accepts particles travelling along a given observation direction $\mathbf{n}$. The active surface element of the detector is therefore $d\mathbf{F}(\mathbf{r}) = |\mathbf{r}|^2 \cdot d\Omega\cdot\mathbf{n}$ and subtends the solid angle $d\Omega$ along $\mathbf{n}$. The detector is set to measure the current of particles crossing the area dF along $\mathbf{n}$, and this current is given by

$$\mathbf{J}(\mathbf{r})\cdot d\mathbf{F}$$

where $\mathbf{J}(\mathbf{r})$ is the probability density vector carried by the wave

$$\psi(\mathbf{r}) = e^{ik\cdot\mathbf{n}_0\cdot\mathbf{x}} + \frac{e^{ik|\mathbf{r}|}}{|\mathbf{r}|}\cdot f(k\cdot(\mathbf{n}_0-\mathbf{n}))$$

$$\mathbf{J} = \frac{\hbar}{i\cdot 2m}\left(\overline{\psi_s(\mathbf{x})}\nabla\psi_s(\mathbf{x}) - \psi_s(\mathbf{x})\nabla\overline{\psi_s(\mathbf{x})}\right)$$

$\mathbf{J}(\mathbf{r})$ amounts to (up to terms which rapidly fluctuate around zero for large distances and therefore averages out to zero when a small but experimentally unavoidable uncertainty of k is allowed)

$$\mathbf{J}(|\mathbf{x}|,k\cdot,\mathbf{n}_0,\mathbf{n}) = \frac{\hbar k}{m}\cdot\mathbf{n}_0 + \frac{|f(k,\mathbf{n}_0,\mathbf{n})|^2}{|\mathbf{x}|^2}\cdot\frac{\mathbf{r}}{|\mathbf{x}|}$$

The current density vector consists of a plane wave component arising from the incident beam. This component is travelling only in the forward direction. It is measured by the detector when it aligned parallel to direction of the incident beam in front of the target (the forward direction). The second component is also detected away from the forward direction and is provided by the scattered wave: the number of particles arriving at the detector per unit time (the scattered probability current) is given by

$$\frac{\hbar k}{m} \mid f(k,\mathbf{n}_0,\mathbf{n}) \mid^2 \cdot d\Omega$$

An uncertain point about this result is that it has been obtained using ψ_s, which is a non-square integrable function. A practical way to get rid of this uncertainty is to divide the scattered probability current by the absolute value of the current density of the incoming plane wave – $J_0 = \hbar\frac{k}{m}$. This last current is also computed with a non-normalizable wave function, so that by the operation of dividing the two quantities one obtains a physically relevant result which is independent of the hypothetical normalization constant. The ratio between the scattered probability current into the solid angle $d\Omega$ and the incoming probability current density is called the differential scattering cross section $d\sigma(k,\mathbf{n}_0,\mathbf{n})$ and amounts to

$$d\sigma(k,\mathbf{n}_0,\mathbf{n}) = \mid f(k,\mathbf{n}_0,\mathbf{n}) \mid^2 \cdot d\Omega$$

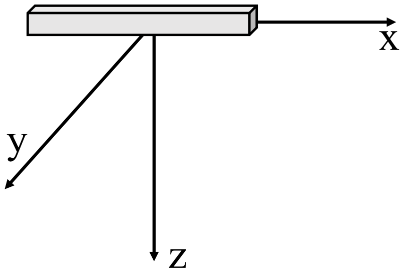
In the literature one defines often the differential cross section as

$$\frac{d\sigma(k,\mathbf{n}_0,\mathbf{n})}{d\Omega} = \mid f(k,\mathbf{n}_0,\mathbf{n}) \mid^2$$

This last result establishes a link between an experimental quantity – the differential cross section – and the scattering amplitude arising from the solution of the SE. We want now to illustrate this result with some simple, computable examples.

EXAMPLES.

1. We consider a potential that is defined on a slit-like region in space: the region has a very small thickness a along z (say, of the order of a lattice constant, i.e. atomically thin), a very small length a along y and a finite size s_0 along x.



In light gray the region of space filled with a constant potential Φ_0.

The plane wave with $\lambda \gg a$ propagates along $+z$, i.e $\mathbf{n}_0 = (0,0,1)$. In this region of space the potential assume a finite value Φ_0, otherwise it is zero. In the first instance, the detector is placed in the xz-plane along $\mathbf{n} = (\sin\theta, 0, \cos\theta)$. The scattering amplitude amounts to

$$-\frac{1}{4\pi}\frac{2m\cdot\Phi_0}{\hbar^2}\int_{-\frac{a}{2}}^{\frac{a}{2}} dz \int_{-\frac{a}{2}}^{\frac{a}{2}} dy \int_{-\frac{s_0}{2}}^{\frac{s_0}{2}} dx \cdot e^{ik\cdot(-x\cdot\sin\theta + z\cdot(1-\cos\theta))}$$

The integrals over y and z can be performed by setting $y = z = 0$ in the integrand and thus give a multiplicative factor of a^2. The only integral left to be computed is the one over x:

$$\int_{-\frac{s_0}{2}}^{\frac{s_0}{2}} dx \cdot e^{-i\cdot k\cdot\sin\theta\cdot x}$$

This integral is elementary and gives

$$-\frac{2}{k\cdot\sin\theta}\cdot\sin[k\sin\theta\frac{s_0}{2}]$$

Accordingly, the scattered differential cross section amounts to

$$4\cdot(\frac{1}{4\pi}\frac{2m\cdot\Phi_0}{\hbar^2})^2\cdot a^4\cdot\frac{\sin^2[k\cdot\sin\theta\cdot\frac{s_0}{2}]}{[k\cdot\sin\theta]^2}$$

The graph of the function

$$\frac{\sin^2[k\cdot\sin\theta\cdot\frac{s_0}{2}]}{[k\cdot\sin\theta]^2}$$

allows to identify a set of angles at which the scattered intensity is vanishing

$$k\sin\theta\frac{s_0}{2}=p\cdot\pi\Leftrightarrow\sin\theta_p=p\frac{\lambda}{s_0};\quad p=\pm1,\pm2,...$$

The scattered current undergoes well-defined sharp peaks at angles in-between. This specific behaviour is referred to as diffraction of particle waves.

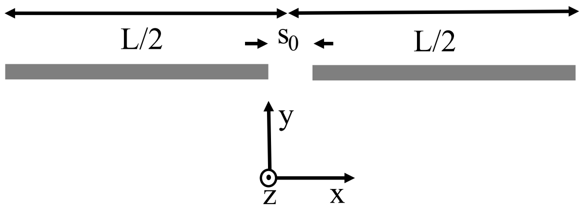
2. We consider the same problem as in the first example. The detector is rotated around the z-axis by 90°, i.e. the observation direction is in the zy-plane and $\mathbf{n}=(0,\sin\theta,\cos\theta)$. The scattering current in this direction is

$$(\frac{1}{4\pi})^2(\frac{2m\Phi_0}{\hbar^2})^2\,|\int_{-\frac{a}{2}}^{\frac{a}{2}}dz\int_{-\frac{a}{2}}^{\frac{a}{2}}dy\int_{-\frac{s_0}{2}}^{\frac{s_0}{2}}dx\cdot e^{ik\cdot(-y\cdot\sin\theta+z\cdot(1-\cos\theta))}\,|^2$$

$$=(\frac{1}{4\pi})^2(\frac{2m\Phi_0}{\hbar^2})^2a^4\cdot s_0^2$$

i.e. it is independent on θ. By detecting the diffracted intensity along two orthogonal directions one would be able to establish that the scattering target has a highly anisotropic shape in the xy-plane and by measuring the angles at which minima occur one would be able to extract the size s_0. Array of detectors covering most the entire solid angle in the hemisphere behind the target are nowadays available for single shot measurement of the diffracted current.

3. Let now the constant potential be active in the two intervals $[-\frac{L}{2},-\frac{s_0}{2}]$ and $[\frac{s_0}{2},\frac{L}{2}]$. Let $L\gg\lambda$. This slit along the x-axis is complementary to the slit used in the previous examples.



In light gray the region of space filled with a constant potential Φ_0.

We are looking for the diffraction pattern in the xz-plane, i.e. $\mathbf{n}=(\sin\theta,0,\cos\theta)$. The scattering amplitude amounts to

$$-\frac{1}{4\pi}\frac{2m\cdot\Phi_0}{\hbar^2}a^2\cdot\underbrace{\left[\int_{-\frac{L}{2}}^{-\frac{s_0}{2}} dx\cdot e^{ik\cdot(-x\cdot\sin\theta)}+\int_{\frac{s_0}{2}}^{\frac{L}{2}} dx\, e^{ik\cdot(-x\cdot\sin\theta)}\right]}_{-\frac{2\cdot\sin(k\sin\theta\frac{s_0}{2})}{k\cdot\sin\theta}+\frac{2\cdot\sin(k\sin\theta\frac{L}{2})}{k\cdot\sin\theta}}$$

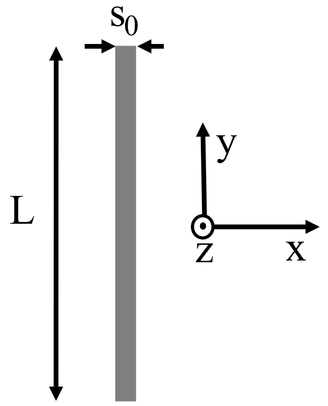
The scattered current amounts to

$$(\frac{1}{4\pi}\frac{2m\cdot\Phi_0}{\hbar^2})^2a^2\cdot$$

$$\left(\frac{4\cdot\sin^2(k\sin\theta\frac{L}{2})}{(k\cdot\sin\theta)^2}+8\cdot\frac{\sin(k\sin\theta\frac{L}{2})}{k\cdot\sin\theta}\frac{\sin(k\sin\theta\frac{s_0}{2})}{k\cdot\sin\theta}+4\cdot\frac{\sin^2(k\sin\theta\frac{s_0}{2})}{(k\cdot\sin\theta)^2}\right)$$

The mixed terms contains a function which is rapidly oscillating between 1 and -1 for large L: the contribution of this term averages to zero because of a small but unavoidably finite angle of aperture of the detector. The first term only contributes a current in the forward direction ($\theta=0$) for very large values of L. The remaining term gives exactly the diffraction pattern computed for the complementary potential. This result is an expression of the famous classical Babinet principle of diffraction (quoting verbatim from Wikipedia): ”In physics, Babinet’s principle states that the diffraction pattern from an opaque body is identical to that from a hole of the same size and shape except for the overall forward beam intensity. It was formulated in the 1800s by the french physicist Jacques Babinet.”

4. The constant potential is defined within a prismatic region with thickness a along z, width s_0 along x and height $L>>\lambda$ along y.



In light gray the region of space filled with a constant potential Φ_0.

The scattering amplitude along $\mathbf{n}=(\sin\theta,0,\cos\theta)$ amounts to

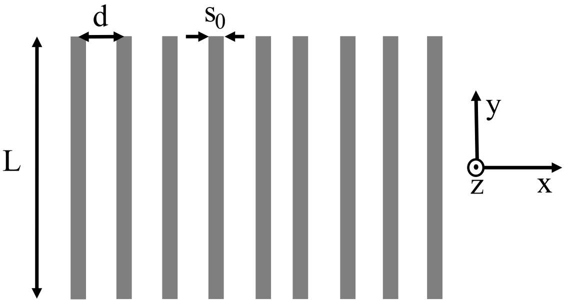
$$\frac{1}{4\pi}\frac{2m\cdot\Phi_0}{\hbar^2}\cdot a\cdot L\cdot\frac{2}{k\cdot\sin\theta}\cdot\sin[k\sin\theta\frac{s_0}{2}]$$

The scattering amplitude along $\mathbf{n}=(0,\sin\theta,\cos\theta)$ amounts to

$$\frac{1}{4\pi}\frac{2m\cdot\Phi_0}{\hbar^2}\cdot a\cdot s_0\cdot\frac{2}{k\cdot\sin\theta}\cdot\sin[k\sin\theta\frac{L}{2}]$$

As $L\gg\lambda$ the scattering amplitude is strongly peaked at $\theta=0$ (i.e. along the z-direction). Because of the length of the slit along y the scattered diffraction pattern evolves with characteristic minima and maxima in the xz-plane, i.e. is concentrated along $y=0$. This is substantially different from the diffraction pattern computed for a slit with length $a \ll \lambda$ in the y-direction.

5. We replicate the prism of the previous example N-times ("grating") in the xy-plane and consider the scattered current along $\mathbf{n} = (\sin\theta, 0, \cos\theta)$.



A sequence of prism along the x-direction with thickeness s_0 and distance d. Along y the slits have a length L.

The periodicity along x-is d. The total scattered amplitude is obtained using the superposition principle:

$$f_N(k,\mathbf{n}_0,\mathbf{n}) = \sum_{j=0,\pm1\ldots}^{\frac{N}{2}} f_j(k,\mathbf{n}_0,\mathbf{n})$$

with

$$f_j(k,\mathbf{n}_0,\mathbf{n}) = -\frac{1}{4\pi}\frac{2m\cdot\Phi_0}{\hbar^2}\cdot a\cdot L\cdot \int_{-\frac{s_0}{2}}^{\frac{s_0}{2}} dx e^{ik\cdot(-(x-j\cdot d))\cdot\sin\theta} = f_0(k,\mathbf{n}_0,\mathbf{n})\cdot e^{-k\cdot j\cdot d\cdot\sin\theta}$$

Accordingly, the total scattered current form the grating of vertical slits amounts to

$$f_0(k,\mathbf{n}_0,\mathbf{n})\cdot \sum_{j=0,\pm1\ldots}^{\frac{N}{2}} e^{-i\cdot k\cdot d\cdot j\cdot\sin\theta}$$

For $N \to \infty$ the $\sum$ is not vanishing only if

$$k\cdot d\cdot\sin\theta = p\cdot 2\pi; \quad p = 0,\pm1,\pm2,\ldots$$

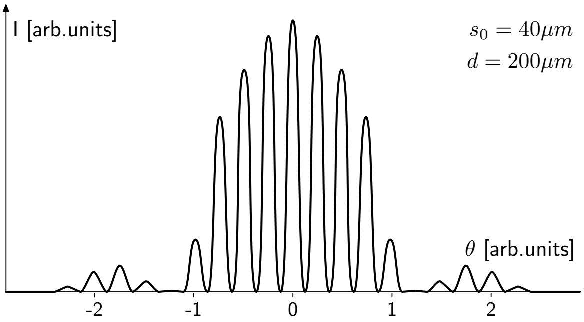
This equation selects a set of observation angles θ_p under which a constructive interference of the diffracted waves takes place, as the sum gives N. These angles are given by the equation

$$\sin\theta_p = \frac{\lambda}{d}\cdot p$$

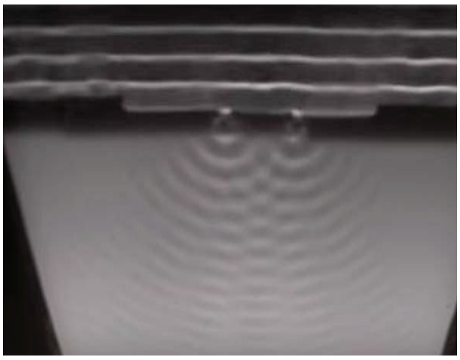
The finite slit width s_0 provides an envelope which modulates the intensity of the sharp intensity maxima by the factor $|\, f_0(k,\mathbf{n}_0,\mathbf{n})\,|^2$. Thus, the diffraction pattern will consist of current maxima with intensity

$$\left(\frac{1}{4\pi}\frac{2m\cdot\Phi_0}{\hbar^2}\cdot a\cdot L\right)^2\cdot N^2\cdot 4\cdot \frac{\sin^2\left[\pi\cdot p\cdot\frac{s_0}{d}\right]}{\left[\pi\cdot p\cdot\frac{s_0}{d}\right]^2}$$

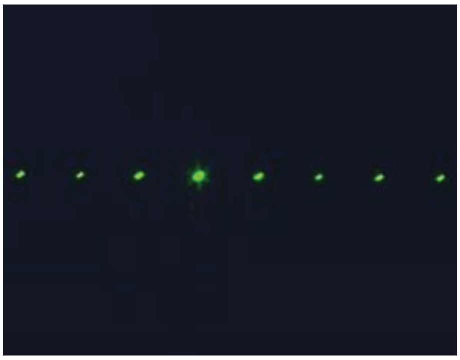
In summary: the interference of the waves from different slits ensures that the diffracted wave is observed only at certain angles. There is no intensity at most angles. In return, the intensity in the maxima is proportional to N^2.



Scattered intensity as a function of the scattering angle θ in the zx-plane. The data refer to a two-slit experiment with $s_0 = 40\mu m$ and $d = 200\mu m$ ($\lambda = 680nm$). The intensity of the peaks is modulated by the single slit scattering intensity as explained in the text.



Photograph of a double slit experiment with water waves.



The diffraction pattern along the x-direction (horizontal in the figure) of green light at a grating established as fine parallel wires in the xy-plane parallel to the y-direction (vertical in the figure). The light is incident along $+z$ (perpendicular to the plane of the paper). `https://experimente.phys.ethz.ch/de/100/10005/20007/30064/`

6. The examples that we have worked out so far can be translated – mutatis mutandis – to the diffraction of any scalar wave at apertures and/or corresponding ”obstacles”. In the specific case one formally construct the inhomogeneous Helmholtz equation which describes a wave encountering a localized region where the refraction index is modified with respect to the one in the remaining region of space:

$$\triangle\psi(\mathbf{r})+k^2\psi(\mathbf{r})=(k^2-k^2(\mathbf{r}))\psi(\mathbf{r})$$

The scattering amplitude from the ”potential”$(k^2-k^2(\mathbf{r}))$ writes

$$f(k\cdot(\mathbf{n}_0-\mathbf{n}))_{\text{Born}}=\int_\Omega d^3y\; e^{ik\cdot(\mathbf{n}_0-\mathbf{n})\cdot\mathbf{y}}\cdot(k^2-k^2(\mathbf{y}))$$

11. Time dependent phenomena in quantum mechanics

11.1. The time dependent Schrödinger equation

In the final paper of 1926, Schrödinger was confronted with the problem of finding an equation for describing the temporal evolution of a quantum mechanical state ψ, given the Helmholtz like appearance of the time-independent wave equation used to find the eigenvalues of the H-atom. Such a temporal evolution is not explicit when dealing with eigenstates of operators, which, by definition, are stationary states of a system. However, a temporal evolution might appear for a given time-independent Hamiltonian when a wave function was determined at a certain time – for instance by localizing a particle in a certain region of space – and then asking into which state the particle would evolve, driven by the Hamilton operator. Or else one could think of a situation where the Hamilton operator contains explicitly a time-dependent component. In both cases one needs an equation that goes beyond the Helmholtz-like wave equation which contains not the time but a fixed energy. The finding of the time dependence Schrödinger equation can be considered as the real element of novelty introduced by Schrödinger: he created an entirely new equation of physics, never seen before – and he did this by ”guessing” the simplest time-dependent equation that, suitably treated, would produce the time-independent equation he used for the Hydrogen atom. Concretely: the classical Helmoltz equation

$$\triangle\psi(\vec{r})+\frac{\omega^2}{v^2}\psi(\vec{r})=0$$

is obtained by inserting the monochromatic wave Ansatz

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

into the classical wave equation

$$\frac{\partial^2\psi(\vec{r},t)}{\partial t^2}=v^2\triangle\psi(\vec{r},t)$$

Schrödinger had his time independent eigenvalue equation

$$\triangle\psi(x,y,z)+\frac{2m}{\hbar^2}\cdot[E-\phi(x,y,z)]\cdot\psi(x,y,z)=0$$

as a starting point. This equation has as solutions the eigenstates of the system with a well defined energy E. Such eigenstates have the property that the probability density does not change in time (convince yourself by building the $|..|^2$). For eigenstates, the time dependent component can only be a prefactor the contains the time as $e^{i\cdot\ldots t}$. In analogy to the classical monochromatic Ansatz Schrödinger uses for eigenstates the stationary states Ansatz

$$\psi(\vec{r},t)=\psi(\vec{r})\cdot e^{-i\frac{E}{\hbar}\cdot t}$$

E being the energy eigenvalue for the sought for eigenstates. The energy of the eigenstate appears in the exponent of the time dependent exponential prefactor $e^{-i\frac{E}{\hbar}\cdot t}$. Now the task is one of guessing the simplest equation that is transformed into the time independent SE after inserting the Ansatz for stationary states. Schrödinger proposed the time dependent Schrödinger equation

$$\boxed{i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t)=-\frac{\hbar^2}{2m}\triangle\psi(\vec{r},t)+\phi(\vec{r})\psi(\vec{r},t)}$$

When one searches for the eigenstates of a system starting from the time dependent SE one must use the stationary states Ansatz $\psi(\vec{r})\cdot e^{-i\frac{E}{\hbar}\cdot t}$. Such states have been the subject of the previous Chapters. But there are other physical situations where the sought for wave-function has a different time dependence. These situations we would like to study in the present Chapter.

COMMENTS.

1. In term of the evolution of quantum mechanical states, this equation can be generalized to any Hamilton operator:

$$i\hbar\frac{d}{dt}\psi=\hat{H}\psi$$

2. Although the time-independent SG has the structure of a Helmholtz equation, the time-dependent SG does not have the structure of a classical wave equation and accordingly the wave function $\psi(\vec{r},t)$ is not a "wave" in the classical sense.

3. Together with the probability density

$$\rho(\vec{r},t)=\overline{\psi(\vec{r},t)}\cdot\psi(\vec{r},t)$$

one can define a probability current density

$$\vec{J}(\vec{r},t)\doteq\frac{\hbar}{2im}\left(\overline{\psi(\vec{r},t)}\vec{\nabla}\psi(\vec{r},t)-(\vec{\nabla}\overline{\psi(\vec{r},t)})\psi(\vec{r},t)\right)$$

which obeys the continuity equation

$$\frac{\partial\rho}{\partial t}+\vec{\nabla}\cdot\vec{J}=0$$

4. By expanding ψ as a superposition of basis states $\{\psi_i\}$ we obtain:

$$i\hbar\frac{d}{dt}\sum_j c_j\psi_j=\hat{H}\sum_j c_j\psi_j$$

We scalar product both sides from the left with ψ_i and use the orthogonality relations for basis functions to obtain a set of ordinary differential equations for the sought for coefficients c_i:

$$i\cdot\hbar\cdot\begin{pmatrix}\dot{c}_1\\ \dot{c}_2\\ \cdot\\ \cdot\\ \dot{c}_n\\ \cdot\end{pmatrix}=\begin{pmatrix}(\psi_1,\hat{H}\psi_1) & (\psi_1,\hat{H}\psi_2) & \cdot & \cdot & \cdot\\ (\psi_2,\hat{H}\psi_1) & (\psi_2,\hat{H}\psi_2) & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ (\psi_n,\hat{H}\psi_1) & (\psi_n,\hat{H}\psi_2) & \cdot & \cdot & \cdot\\ \cdot & & & & \end{pmatrix}\begin{pmatrix}c_1\\ c_2\\ \cdot\\ \cdot\\ c_n\\ \cdot\end{pmatrix}$$

(shortly:

$$i\cdot\hbar\cdot\dot{c}_i=\sum_j\left(\psi_i,\hat{H}\psi_j\right)c_j$$

11.2. Simple two-states problems

In this section we investigate some two-states systems subject to time-dependent phenomena. These are mathematically speaking most elementary systems where time-dependent phenomena, relevant in modern research, can be treated without loosing track of the physical relevance. We consider the time dependent Schrödinger equation of a spin $\frac{1}{2}$ system in a uniform magnetic field:

$$i\hbar\begin{pmatrix}\dot{C}^+\\ \dot{C}^-\end{pmatrix}=\mu_B\cdot\begin{pmatrix}B_z & B_x-i\cdot B_y\\ B_x+i\cdot B_y & -B_z\end{pmatrix}\begin{pmatrix}C^+\\ C^-\end{pmatrix}$$

(when spinors are involved, the SE is often referred to as the Pauli equation). The sought for coefficients describe the state $C^+(t)\cdot Y^+ + C^-(t)\cdot Y^-$, i.e. when the coefficients are found one knows that the state that solve the Pauli equation is a superposition of "up" and "down" z-spin components with the amplitudes $C^+(t)$ and $C^-(t)$.

Zeeman splitting We consider as a first example the simple case where $B_x = B_y = 0$, i.e. we set the constant magnetic field along the z-axis. We look first for the stationary states of the Pauli equation. After inserting the Ansatz $C^{\pm}(t) = e^{\mp\frac{i}{\hbar}\cdot E\cdot t}$ we obtain the time independent eigenvalue equation

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \mu_B\cdot B_z - E & 0 \\ 0 & -\mu_B\cdot B_z - E \end{pmatrix} \begin{pmatrix} C^+ \\ C^- \end{pmatrix}$$

This is a set of algebraic equations. The determinantal equation

$$(\mu_B\cdot B_z - E)\cdot(-\mu_B\cdot B_z - E) = 0$$

has the two solutions. One is $E_+ = \mu_B\cdot B_z$. With this energy, the system of linear equation is solved by the eigenstate $(1,0)$, which corresponds to the stationary state $e^{-i\mu_B\cdot B_z\cdot t}\cdot Y^+$. The second one is $E_+ = -\mu_B\cdot B_z$. With this energy, the system of linear equation is solved by the eigenstate $(0,1)$, which corresponds to the stationary state $e^{-i(-\mu_B\cdot B_z)\cdot t}\cdot Y^-$. These two states correspond to the Zeeman splitted states, with splitting $2\cdot\mu_B\cdot B$. This result shows that the ground state of this two-state system consists of "down" spins and that the z component of the magnetic moment is, correspondingly, $+\mu_B$.

Initial conditions. We consider again the situation $B_x = B_y = 0$ but from a different point of view. We ask: how can we deposit a spin in one of the stationary states? Let us assume that we have a magnetic field along z and we have succeeded in placing the spin along $-z$ at some time, which we choose to be $t = 0$. We seek now a solution of the Pauli equation with a given initial condition

$$\begin{pmatrix} C_+(t=0) = 0 \\ C_-(t=0) = 1 \end{pmatrix}$$

which describes a spin residing in the state of energy $-\mu_B\cdot B$ at $t = 0$. The general solution of the two differential equations can be found to be:

$$\begin{pmatrix} C_+(t) = a\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}t} \\ C_-(t) = b\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}t} \end{pmatrix}$$

with a and b being integration constants. The initial condition can be fulfilled by taking

$$\begin{pmatrix} a = 0 \\ b = 1 \end{pmatrix}$$

meaning that the solution to this initial conditions is

$$\begin{pmatrix} C_+(t) = 0 \\ C_-(t) = e^{-i\frac{\mu_B\cdot B}{\hbar}t} \end{pmatrix}$$

As

$$|C_-(t)|^2 = 1$$

at any time, this solution means that the spin remains in the state of energy $-\mu_B\cdot B$ at any time. So, by choosing the suitable initial condition we have succeded in depositing the spin into one of the stationary states. Of course, the next question is: how did we get the spin along $-z$ at $t = 0$, i.e. how did we obtain the initial condition? The answer to this question is a really difficult experimental protocol which goes beyond the scope of this lecture.

Spin precession. Much more interesting is the situation where the initial condition is the same but a magnetic field of strength B_1 is applied along, e.g. the x-direction. The Pauli equation writes

$$i\hbar\dot{C}^{+} = \mu_B \cdot B_1 \cdot C^{-}$$
$$i\hbar\dot{C}^{-} = \mu_B \cdot B_1 \cdot C^{+}$$

To find the general solution we add and subtract the two equations, which then are decoupled to equations for the new variables $A^{\pm} \doteq \dot{C}^{+} \pm \dot{C}^{-}$:

$$i\hbar\dot{A}^{+} = \mu_B \cdot B_1 \cdot A^{+}$$
$$i\hbar\dot{A}^{-} = -\mu_B \cdot B_1 \cdot A^{-}$$

with general solutions

$$A^{+} = a \cdot e^{-\frac{i}{\hbar}\mu_B \cdot B_1 \cdot t}$$
$$A^{-} = b \cdot e^{+\frac{i}{\hbar}\mu_B \cdot B_1 \cdot t}$$

These general solution can be translated to general solutions for $C^{\pm}$:

$$C^{+} = a \cdot e^{-\frac{i}{\hbar}\mu_B \cdot B_1 \cdot t} + b \cdot e^{+\frac{i}{\hbar}\mu_B \cdot B_1 \cdot t}$$
$$C^{-} = a \cdot e^{-\frac{i}{\hbar}\mu_B \cdot B_1 \cdot t} - b \cdot e^{+\frac{i}{\hbar}\mu_B \cdot B_1 \cdot t}$$

We seek now to find the special solution that fulfills the initial conditions

$$\begin{pmatrix} C_{+}(t=0) = 0 \\ C_{-}(t=0) = 1 \end{pmatrix}$$

The initial condition specifies the integration constants a and b to be $-b = a = \frac{1}{2}$, so that we find the solution to the initial condition to be

$$C^{+} = -i \cdot \sin \tfrac{\mu_B \cdot B_1}{\hbar} \cdot t$$
$$C^{-} = +1 \cdot \cos \tfrac{\mu_B \cdot B_1}{\hbar} \cdot t$$

These coefficients define a time dependent state

$$\psi(t) = -i \sin(\frac{\mu_B \cdot B_1}{\hbar} \cdot t) \cdot Y^{+} + \cos(\frac{\mu_B \cdot B_1}{\hbar} \cdot t) \cdot Y^{-}$$

The probability that the spin is found in the "up"-state amounts to

$$\mid C^{+} \mid^2 = \sin^2 \frac{\mu_B \cdot B_1}{\hbar} \cdot t$$

In a static transversal magnetic field there is a finite probability for spin-flip, the probability varying periodically with a frequency $\frac{\mu_B \cdot B_1}{\hbar}$. If we think of an ensemble of spin polarized electrons entering a region of transversal magnetic field, we can compute the spin polarization vector of the electrons, defined as

$$< \vec{P} >|_{\psi(t)} \doteq \left(< \sigma_x >|_{\psi(t)}, < \sigma_y >|_{\psi(t)}, < \sigma_z >|_{\psi(t)}\right)$$

In this situation, we have

$$< \sigma_x >|_{\psi(t)} = \begin{pmatrix} -i \sin \frac{\mu_B \cdot B_1}{\hbar} \cdot t & \cos \frac{\mu_B \cdot B_1}{\hbar} \cdot t \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -i \sin \frac{\mu_B \cdot B_1}{\hbar} \cdot t \\ \cos \frac{\mu_B \cdot B_1}{\hbar} \cdot t \end{pmatrix} = 0$$

$$< \sigma_y >|_{\psi(t)} = \begin{pmatrix} -i \sin \frac{\mu_B \cdot B_1}{\hbar} \cdot t & \cos \frac{\mu_B \cdot B_1}{\hbar} \cdot t \end{pmatrix} \cdot \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} -i \sin \frac{\mu_B \cdot B_1}{\hbar} \cdot t \\ \cos \frac{\mu_B \cdot B_1}{\hbar} \cdot t \end{pmatrix} =$$

$$2 \cdot \sin(\frac{\mu_B \cdot B_1}{\hbar} \cdot t) \cdot \cos(\frac{\mu_B \cdot B_1}{\hbar} \cdot t) = \sin(2 \cdot \frac{\mu_B \cdot B_1}{\hbar} \cdot t)$$

$$< \sigma_z >|_{\psi(t)} = \begin{pmatrix} -i \sin \frac{\mu_B \cdot B_1}{\hbar} \cdot t & \cos \frac{\mu_B \cdot B_1}{\hbar} \cdot t \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -i \sin \frac{\mu_B \cdot B_1}{\hbar} \cdot t \\ \cos \frac{\mu_B \cdot B_1}{\hbar} \cdot t \end{pmatrix} =$$

$$\sin^2(\frac{\mu_B \cdot B_1}{\hbar} \cdot t) - \cos^2(\frac{\mu_B \cdot B_1}{\hbar} \cdot t) = -\cos(2 \cdot \frac{\mu_B \cdot B_1}{\hbar} \cdot t)$$

This vector has the property that its length is 1, independent of time. It rotates in a plane perpendicular to the axis defined by the direction of the magnetic field with the frequency

$$2\cdot\frac{\mu_B\cdot B_1}{\hbar}$$

This motion of the spin is called "precessional motion". Excitation of precessional motion plays in many applications an important role.

11.3. *Magnetic resonance

(I. Rabi, Phys. Rev. **51**, 652 (1937), Nobel Preis 1944). We now solve the problem originally solved by I. Rabi and implemented in a famous experiment demonstrating magnetic resonance. A strong, constant in time magnetic field

$$(0,0,B_0)$$

is applied along the z-axis, defining Zeeman split states in a system containing atomic magnetic moments. Transversal to it a small time-dependent magnetic field $\vec{B}_1$ is applied, which, for mathematical simplicity, we take to be

$$\vec{B}_1=(B_1\cdot\cos\omega t, B_1\cdot\sin\omega t, 0)$$

The Pauli equation for a spin $\frac{1}{2}$ system reads:

$$i\hbar\begin{pmatrix}\dot{C}_+\\ \dot{C}_-\end{pmatrix}=\left(\begin{array}{c|c}B_0\mu_b & \mu_B B_1 e^{-i\omega t}\\ \hline \mu_B B_1 e^{i\omega t} & -B_0\mu_B\end{array}\right)\begin{pmatrix}C_+\\ C_-\end{pmatrix}$$

This is a set of 1. order (in t) coupled differential equations, with coefficients depending on t explicitly. Given the initial condition

$$\begin{pmatrix}C_+(t=0)=0\\ C_-(t=0)=1\end{pmatrix}$$

one obtains the famous Rabi formula as solution:

$$|C_+|^2=\frac{4\omega_1^2}{(\omega_L-\omega)^2+4\omega_1^2}\cdot\sin^2(\frac{\Omega}{2}t)$$

Those interested at the mathematical details of the solution are referred to the end of this section. The various characteristic frequencies entering this formula are related to the original parameters entering the Hamilton operator by the following relations:

$$\Delta=E_+-E_-$$

is the Zeeman splitting energy.

$$\omega_L=\frac{\Delta}{\hbar}$$

$$\omega_1=\frac{\mu_B\cdot B_1}{\hbar}$$

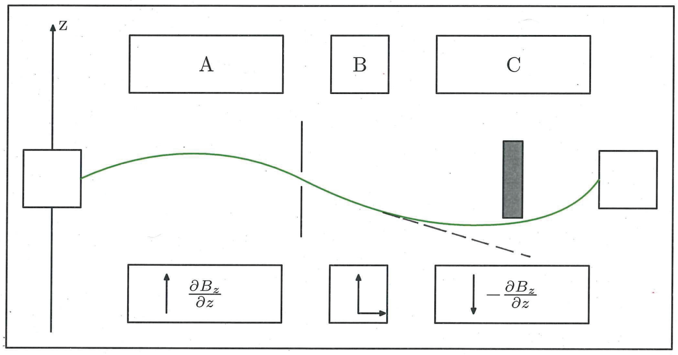
$$\Omega=\sqrt{(\omega_L-\omega)^2+4\omega_1^2}$$

In the experimental situation forthcoming in the Rabi experiment (see below), the time at disposal for the transition is small but there is a finite probability for spin-down to make the transition to the spin up state, which depends on the frequency ω by virtue of the function

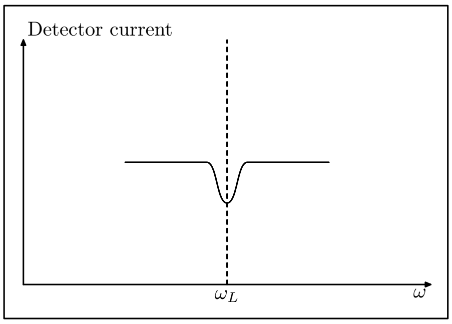
$$\frac{4\omega_1^2}{(\omega_L-\omega)^2+4\omega_1^2}$$

A plot of this function versus ω – the frequency of the transversal magnetic field $\vec{B}_1$ – gives the famous Lorenz curve centered at ω_L, which unequivocally defines this quantum mechanical transition as resonance process. ω_L is the so-called Larmor frequency at which the transition probability versus frequency acquires a sharp resonance peak (the smallest ω_1 the sharpest is the peak).

For the practical application of this formula, Rabi used a modified Stern-Gerlach device.



The Rabi experiment. An "oven" produces atoms (or any other particle such as neutrons with a net total angular momentum). Upon entering the magnet A, they are sorted out by the gradient of the magnetic field in separate beams according to the z component of the angular momentum. Magnet A thus establishes the initial condition under which the atoms enters magnet B. In magnet B there is an homogeneous magnetic field B_0 *and a homogeneous oscillating magnetic field* B_1 *perpendicular to* B_0*. If* $B_1 = 0$*, then the particles leave magnet B in the same eigenstate they entered and in magnet C they will be turned back to a position along the axis (continuos green line): almost all atoms will arrive at the detector (right). When* B_1 *is switched on, some atoms will change their spin state in the region of magnet B and follow the dashed trajectory: the detector will count a lesser number of atoms.*



Typical curve observed in a Rabi experiment. When the frequency ω *of the oscillating magnetic field is varied, the number of atoms detected at the counter will follow the function* $\frac{4\omega_1^2}{(\omega_L-\omega)^2+4\omega_1^2}$ *and a typical resonance signal will be observed, from which we can read out, very precisely, the resonance frequency* ω_L*, which gives us information about the g-factor of the particle. Such magnetic resonance experiments are used to determine accurately the magnetic moment of particles.*

Mathematical Details of the Rabi solution

Using the characteristic frequencies

$$\frac{B_0\mu_B}{\hbar} = \omega_0$$

and

$$\frac{B_1\mu_B}{\hbar}=\omega_1$$

we obtain

$$i\begin{pmatrix} \dot{C}_1 \\ \dot{C}_- \end{pmatrix}=\begin{pmatrix} \omega_0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & -\omega_0 \end{pmatrix}\begin{pmatrix} C_+ \\ C_- \end{pmatrix}$$

We insert the Ansatz

$$\begin{pmatrix} C_+ \\ C_- \end{pmatrix}=\begin{pmatrix} e^{-i\omega_0 t}F(t) \\ e^{i\omega_0 t}G(t) \end{pmatrix}$$

and obtain

$$\begin{aligned} i\left(-i\omega_0 e^{-i\omega_0 t}F(t)+e^{-i\omega_0 t}\dot{F}(t)\right) &= \omega_0 F(t)e^{-i\omega_0 t}+\omega_1 e^{-i\omega t}G(t)e^{i\omega_0 t} \\ i\left(i\omega_0 e^{i\omega_0 t}G(t)+e^{i\omega_0 t}\dot{G}\right) &= \omega_1 e^{i\omega t}F(t)e^{-i\omega_0 t}-\omega_0 G e^{i\omega_0 t} \end{aligned}$$

$$\left\{ \begin{array}{lcl|l} i\dot{F}(t)e^{-i\omega_0 t} & = & \omega_1 e^{-i\omega t}G(t)e^{i\omega_0 t} & e^{i\omega_0 t} \\ i\dot{G}(t)e^{i\omega_0 t} & = & \omega_1 e^{i\omega t}F(t)e^{-i\omega_0 t} & e^{-i\omega_0 t} \end{array}\right.$$

$$i\begin{pmatrix} \dot{F} \\ \dot{G} \end{pmatrix}=\omega_1\begin{pmatrix} G e^{i(2\omega_0-\omega)t} \\ F e^{-i(2\omega_0-\omega)t} \end{pmatrix}$$

We define further

$$2\omega_0=\frac{\Delta}{\hbar}\doteq\omega_L$$

as the so-called Larmor frequency. Then

$$i\begin{pmatrix} \dot{F} \\ \dot{G} \end{pmatrix}=\omega_1\begin{pmatrix} G e^{i(\omega_L-\omega)t} \\ F e^{-i(\omega_L-\omega)t} \end{pmatrix}$$

The coupled system can be decoupled by taking the time derivative:

$$\begin{aligned} i\ddot{F} &= \omega_1 G i(\omega_L-\omega)e^{i(\omega_L-\omega)t}+\omega_1\dot{G}e^{i(\omega_l-\omega)t} \quad |\cdot i \\ &\Rightarrow \quad -\ddot{F}=-\omega_1(\omega_L-\omega)G e^{i(\omega_L-\omega)t}+\omega_1^2 F \\ &\Rightarrow \quad -\ddot{F}=-i(\omega_L-\omega)\dot{F}+\omega_1^2 F \\ &\Rightarrow \quad \ddot{F}-i(\omega_L-\omega)\dot{F}+\omega_1^2 F=0 \end{aligned}$$

For finding the general solution we use the Ansatz $F=e^{ip}$ ⇒ Then

$$-p^2+p(\omega_L-\omega)+\omega_1^2=0$$

$$p_{\mp}=\frac{(\omega_L-\omega)\mp\sqrt{(\omega_L-\omega)^2+4\omega_1^2}}{2}$$

$$\sqrt{(\omega_L-\omega)^2+4\omega_1^2}=\Omega$$

$$F=a\cdot e^{i(\frac{\omega_L-\omega-\Omega}{2})t}+b\cdot e^{i(\frac{\omega_L-\omega+\Omega}{2})t}$$

a, b are integration constants that will be determined by the initial conditions. For the function G we find

$$\begin{aligned} G(t)= \quad & \tfrac{a\cdot 2\omega_1}{\omega_L-\omega+\Omega}e^{i(\omega_L-\omega+\Omega)t} \\ & +\tfrac{b\cdot 2\omega_1}{\omega_L-\omega-\Omega}e^{-i(\omega_L-\omega-\Omega)t} \end{aligned}$$

Finally, the general solution reads

$$\begin{aligned} C_+(t) &= e^{-i\omega_L/2t}\left(a e^{i(\frac{\omega_L-\omega-\Omega}{2})t}+b e^{i(\frac{\omega_L-\omega+\Omega)}{2}t}\right) \\ C_-(t) &= e^{i\omega_L/2t}\left(a\frac{2\omega_1}{(\omega_L-\omega+\Omega)}e^{i(\frac{\omega_L-\omega+\Omega}{2})t}+b\frac{2\omega_1}{\omega_L-\omega-\Omega}e^{-i(\frac{\omega_L-\omega-\Omega)}{2}t}\right) \end{aligned}$$

With the initial conditions

$$\left\{ \begin{array}{lcl} C_+(t=0) & = & 0 \\ C_-(t=0) & = & 1 \end{array}\right\}$$

we have

$$\left\{ \begin{array}{ll} a+b=0 & \\ \frac{a2\omega_1}{\omega_L-\omega+\Omega}+\frac{b2\omega_1}{(\omega_L-\omega+\Omega)} & =1 \end{array}\right.$$

$$\begin{cases} b = & \frac{(\omega_L-\omega)^2-\Omega^2}{4\Omega\omega_1} = \frac{(\omega_L-\omega)^2-(\omega_L-\omega)^2-4\omega_1^2}{4\sqrt{(\omega_L-\omega)^2+4\omega_1^2}\omega_1} = \frac{-\omega_1}{\sqrt{(\omega_L-\omega)^2+4\omega_1^2}} \\ a = & \frac{\Omega^2-(\omega_L-\omega)^2}{4\Omega\omega_1} = \frac{\omega_1}{\sqrt{(\omega_L-\omega)^2+4\omega_1^2}} \end{cases}$$

Computing explicitely the probability of being in a state Y^+ we obtain

$$C_+(t) = e^{-i\omega_L/2t}\left[\frac{\omega_1}{\sqrt{(\omega_L-\omega)^2+4\omega_1^2}}e^{i\frac{(\omega_L-\omega-\Omega)}{2}t} - \frac{\omega_1}{\sqrt{(\omega_L-\omega)^2+4\omega_1^2}}e^{i\frac{(\omega_L-\omega+\Omega)}{2}t}\right]$$

$$= e^{-i\frac{\omega_L t}{2}}e^{i\frac{(\omega_L-\omega)}{2}t}\left[\frac{\omega_1}{\sqrt{\cdots}}e^{-i\frac{\Omega}{2}t} - \frac{\omega_1}{\sqrt{\cdots}}e^{i\frac{\Omega}{2}t}\right]$$

$$= e^{-i\frac{\omega}{2}t}\frac{\omega_1}{\sqrt{\cdots}}\left[-i\sin\frac{\Omega}{2}t - i\sin\frac{\Omega}{2}t\right] = e^{-i\frac{\omega}{2}t}\frac{-2i\omega_1}{\sqrt{\cdots}}\sin\frac{\Omega}{2}t$$

and the famous Rabi formula

$$\mid C_+\mid^2 = \frac{4\omega_1^2}{(\omega_L-\omega)^2+4\omega_1^2}\cdot\sin^2(\frac{\Omega}{2}t)$$

11.4. * Time-dependent Hamiltonian: perturbational approach

We consider a situation where a system is, at a time $-T$, in a stationary state $e^{-\frac{E_i}{\hbar}\cdot t}\cdot\psi_i$ of the operator H_0. Between $-T$ and $+T$ the system is perturbed by an operator H_1 which can depend on the time t. The operator $\hat{H}_0+\hat{H}_1(t)$ translates the state

$$e^{-\frac{E_i}{\hbar}\cdot t}\cdot\psi_i \quad \text{to} \quad \psi(T)$$

according to the time-dependent Schrödinger equation. Given a state

$$e^{-\frac{E_f}{\hbar}\cdot t}\cdot\psi_f \quad \text{with} \quad E_f \neq E_i$$

we compute now the probability that the system is in the state ψ_f at the time T, i.e.

$$P_{i\to f} = \mid\left(\psi_f,\psi(T)\right)\mid^2$$

If we consider the CONS $\{\psi_k\}$ of the eigenstates of H_0 and write

$$\psi(T) = \sum_k c_k(t)\cdot\psi_k$$

than the sought for probability is given by $\mid c_f(t)\mid^2$. The time-dependent Schrödinger equation translates into an exact equation for then sought for coefficient $c_f(t)$

$$i\hbar\dot{c}_f = \left(\psi_f,H_0+H_1(t)\psi_1\right)c_1+\ldots+\left(\psi_f,H_0+H_1(t)\psi_i\right)c_i+\left(\psi_f,H_0+H_1(t)\psi_f\right)c_f+\ldots.$$

(for simplicity, we have ordered the states so that $f=i+1$). This equation can be simplified factorizing the time-dependence that would arise without perturbation and writing

$$c_k(t) = C_k(t)e^{-i\frac{E_k}{\hbar}\cdot t}$$

Inserting this Ansatz and considering the orthogonality properties of the basis functions we obtain an equation for the coefficient $C_f(t)$:

$$i\hbar e^{-i\frac{E_f}{\hbar}\cdot t}\dot{C}_f = \left(\psi_f,H_1(t)\psi_1\right)e^{-i\frac{E_1}{\hbar}\cdot t}C_1+\ldots+\left(\psi_f,H_1(t)\psi_i\right)e^{-\frac{E_i}{\hbar}\cdot t}C_i+\ldots..C_{j\neq i}\ldots$$

This equation is again formally exact but cannot be solved exactly. However, we can transform it into a solvable equation by the following approximation:

- The system at -T is such that $C_i = 1$, all other coefficients are vanishing.
- At times larger than T, the coefficients $C_{j\neq i}$ are assumed to be small and $C_i \approx 1$.

In order to render the differential equation for $C_f(t)$ solvable, we insert $C_i = 1$ and $C_{j\neq i} = 0$ in the right hand side of the differential equation, which becomes

$$\dot{C}_f(t) = \left(\psi_f, H_1(t)\psi_i\right) \cdot e^{\frac{i}{\hbar}(E_f - E_i)\cdot t}$$

and has the solution

$$C_f(T) = \frac{1}{i\cdot\hbar} \cdot \int_{-T}^{T} dt \left(\psi_f, H_1(t)\psi_i\right) \cdot e^{\frac{i}{\hbar}(E_f - E_i)\cdot t}$$

The assumption underlining this approximation is that $\mid C_f(T) \mid << 1$. To render this assumption in terms of the parameters ruling the transition $i \rightarrow f$ we define the largest value that the matrix element can take in the time interval $2T$ as

$$\mid \left(\psi_f, H_1(t)\psi_i\right) \mid_{max}$$

and obtain the inequality

$$\mid C_f(T) \mid \leq \frac{1}{\hbar} \cdot 2 \cdot T \cdot \mid \left(\psi_f, H_1(t)\psi_i\right) \mid_{max} << 1$$

Further computations of the coefficient depend on the nature of the perturbation operator $H_1(t)$.

11.4.1. Monochromatic periodic perturbation

An important situation is provided by a periodic perturbation

$$H_1(t) = H_1 \cdot e^{-i\omega t} + H_1^{\dagger} \cdot e^{i\omega t}$$

H_1 shall depend only on $\vec{r}$ and m_s. For this situation the transition amplitude writes

$$C_f(T) = \frac{1}{i\hbar} \cdot [\int_{-T}^{T} dt \cdot e^{\frac{i}{\hbar}(E_f - E_i - \hbar\omega)\cdot t} \left(\psi_f, H_1\psi_i\right) + \int_{-T}^{T} dt \cdot e^{\frac{i}{\hbar}(E_f - E_i + \hbar\omega)\cdot t} \left(\psi_f, H_1^{\dagger}\psi_i\right)$$

The integral over the variable t gives

$$\frac{2 \cdot \sin \frac{1}{\hbar}(E_f - E_i - \hbar\omega) \cdot T}{\frac{i}{\hbar}(E_f - E_i - \hbar\omega)}$$

(left-hand side) respectively

$$\frac{2 \cdot \sin \frac{1}{\hbar}(E_f - E_i + \hbar\omega) \cdot T}{\frac{i}{\hbar}(E_f - E_i + \hbar\omega)}$$

It is interesting to plot the graphs of these two functions for a fixed time T as a function of the energy $\hbar\omega$. The graph on the left-hand side falls abruptly to zero when $\hbar\omega \neq E_f - E_i$ and oscillates with smaller amplitude around zero. The graph on the right-hand side, instead, falls abruptly to zero when $\hbar\omega \neq E_i - E_f$ and oscillates with smaller amplitude around zero. In both cases, the amplitude is concentrated within a width $\hbar\Delta\omega \approx \frac{2\pi\hbar}{T}$, meaning that, if T is sufficiently large, the two graphs are well separated.

We apply these results to a two-states system with levels $E_2 > E_1$

1. Let the initial state i be the one with energy E_1. Then the amplitude for transition into the higher lying state reads

$$C_{1\rightarrow 2}(T) = \frac{1}{i\hbar}(\psi_2, H_1\psi_1) \cdot 2 \cdot \frac{\sin \frac{1}{\hbar}(E_2 - E_1 - \hbar\omega) \cdot T}{\frac{i}{\hbar}(E_2 - E_1 - \hbar\omega)}$$

and

$$\mid C_{1\rightarrow 2}(T) \mid^2 = \frac{1}{\hbar^2} \mid (\psi_2, H_1\psi_1) \mid^2 \mid \frac{2 \cdot \sin \frac{1}{\hbar}(E_2 - E_1 - \hbar\omega) \cdot T}{\frac{i}{\hbar}(E_2 - E_1 - \hbar\omega)} \mid^2$$

We evaluate

$$|\frac{2\cdot\sin\frac{1}{\hbar}(E_2-E_1-\hbar\omega)\cdot T}{\frac{i}{\hbar}(E_2-E_1-\hbar\omega)}|^2$$

by writing it has

$$\pi\cdot 4\cdot T\cdot\frac{\sin^2\frac{1}{\hbar}(E_2-E_1-\hbar\omega)\cdot T}{\frac{1}{\hbar^2}(E_2-E_1-\hbar\omega)^2\cdot\pi\cdot T}$$

and considering that

$$\frac{\sin^2\frac{1}{\hbar}(E_2-E_1-\hbar\omega)\cdot T}{\frac{1}{\hbar^2}(E_2-E_1-\hbar\omega)^2\cdot\pi\cdot T}$$

resembles very much

$$\hbar\delta(E_2-E_1-\hbar\omega)$$

for sufficiently long times T. This result can then be formally cast into a transition rate $\Gamma_{1\to 2}\doteq\frac{|C_{1\to 2}(T)|^2}{2T}$ with

$$\Gamma_{1\to 2}=\frac{2\pi}{\hbar}\,|(\psi_2,H_1\psi_1)|^2\cdot\delta(E_2-E_1-\hbar\omega)$$

which is so important that it is called the "Fermi golden rule". This formula explains that the transition from an atomic state 1 to an atomic state 2 with higher energy occurs with highest probabilty when the frequency of the perturbation matches the energy difference between the states (the Bohr frequency condition). The relevant matrix elements to be computed is

$$(\psi_2,H_1\psi_1)$$

2. Let the initial state be the one with energy E_2. In this case the amplitude writes

$$C_{2\to 1}(T)=\frac{1}{i\hbar}\left(\psi_1,H_1^\dagger\psi_2\right)\cdot 2\cdot\frac{\sin\frac{1}{\hbar}(E_1-E_2+\hbar\omega)\cdot T}{\frac{i}{\hbar}(E_1-E_2+\hbar\omega)}$$

Using similar technical arguments we can work out the transition rate from a higher energy states into a lower lying one:

$$\Gamma_{2\to 1}=\frac{2\pi}{\hbar}\,|\left(\psi_1,H_1^\dagger\psi_2\right)|^2\cdot\delta(E_2-E_1-\hbar\omega)$$

In this case the relevant matrix element is

$$\left(\psi_1,H_1^\dagger\psi_2\right)=\overline{(\psi_2,H_1\psi_1)}$$

We find the rather important result for the transition rates:

$$\Gamma_{1\to 2}=\Gamma_{2\to 1}$$

which proves the original idea by Einstein that the absorption and emission coefficients by transitions are identical.

Optical transitions in an atom.

We apply these results to compute optical transition rates in atomic systems. The Hamiltonian that describes the coupling of an electron to a radiation field is given by

$$\frac{1}{2m}\cdot(\vec{p}-q\vec{A})^2+\phi(\vec{r})$$

We recognize

$$H_0=\frac{\vec{p}^2}{2m}+\phi(\vec{r})$$

and

$$H_1(t)=-\frac{q}{2m}(\vec{p}\cdot\vec{A}+\vec{A}\cdot\vec{p})=-\frac{q}{2m}(\underbrace{(\vec{p}\cdot\vec{A})}_{\sim\vec{\nabla}\vec{A}=0}+2\cdot\vec{A}\cdot\vec{p})=-\frac{q}{m}\cdot\vec{A}(\vec{r},t)\cdot\vec{p}$$

We now use the relation

$$\frac{i}{\hbar}[H_0,\vec{r}]=\frac{\vec{p}}{m}$$

to find an expression for transition operator that is found often in the literature. Write first

$$H_1(t)=-q\cdot\frac{i}{\hbar}\vec{A}\cdot[H_0,\vec{r}]$$

The dipolar approximation foresees that the wavelength of the radiation is much longer than the dimensions of an atom, so that $\vec{A}\approx\vec{A}(\vec{r}=0,t)$, the atom residing at the coordinate $\vec{r}=0$. For the optical transition in dipolar approximation between the state i into the state f the coefficient $C_f(t)$ writes

$$C_f(T)=-\frac{1}{i\hbar}\int_{-T}^{T}dt e^{\frac{i}{\hbar}(E_f-E_i)\cdot t}\cdot(q\cdot\frac{i}{\hbar}\vec{A}(t))\cdot\left(\psi_f,[H_0,\vec{r}]\psi_i\right)$$

With

$$q\frac{i}{\hbar}\left(\psi_f,[H_0,\vec{r}]\psi_i\right)=\frac{i}{\hbar}(E_f-E_i)\cdot\left(\psi_f,q\cdot\vec{r}\psi_i\right)$$

and after partial integration over the time variable, we obtain (using $\vec{E}=-\frac{\partial\vec{A}}{\partial t}$)

$$C_f(T)=\frac{1}{i\hbar}\int_{-T}^{T}dt\cdot e^{\frac{i}{\hbar}(E_f-E_i)\cdot t}\cdot\left(\psi_f,(-\vec{E}(t)\cdot\vec{D})\psi_i\right)$$

with $\vec{D}\doteq q\vec{r}$. This is the operator of the electrical dipole moment. We have therefore found an interesting correspondence between the classical situation of the energy of an electric dipole in an electrostatic field, which we know to be $-\vec{E}\cdot\vec{D}$, and the perturbation operator relevant for optical transitions, which we have just found to be

$$H_1(t)=-\vec{E}(t)\cdot\vec{D}$$

Assuming that we can write

$$\vec{E}(t)=\vec{E}_0(\omega)e^{-i\omega t}+\overline{\vec{E}_0}e^{i\omega t}$$

the relevant optical transition matrix elements is therefore

$$\left(\psi_f,-\vec{E}_0\cdot\vec{D}\,\psi_i\right)$$

Let us evaluate this matrix element for some special but important situations

1. We choose monochromatic linearly polarised light propagating along some direction $\vec{k}$. The electric field of the radiation at the site of the atom reads
$$\vec{E}(t)=E_0\cdot\vec{e}_0\cdot\cos(\omega\cdot t)=\frac{\vec{e}_0}{2}\cdot E_0\cdot[e^{-i\cdot\omega\cdot t}+e^{i\cdot\omega\cdot t}]$$
with $\vec{e}_0$ being a unit vector pointing along the polarization direction. The matrix element for optical transition is therefore
$$-q\cdot E_0\cdot\left(\psi_f,\hat{\vec{r}}\cdot\vec{e}_0\psi_i\right)$$
For a system with spherically symmetric field, initial and final states are given by spherical harmonics:
$$\psi_i=f_i(r)\cdot Y_{l_i}^{m_i}(\vartheta,\varphi)$$
$$\psi_f=f_f(r)\cdot Y_{l_f}^{m_f}(\vartheta,\varphi)$$
The spin state does not change with electric dipole transitions, so we do not need to consider the spin state. Because of the presence of the spherical functions, conditions can generally be established for which the matrix elements are found to be exactly vanishing. These conditions are called selection rules. For example, we choose $\vec{e}_0$ in the z direction. Then
$$\vec{r}\cdot\vec{e}_0=z=r\sqrt{\frac{4\pi}{3}Y_1^0(\vartheta,\varphi)}$$

Inserted into the matrix element we obtain

$$\left(\psi_f, \hat{\vec{r}} \cdot \vec{e}_0 \psi_i\right) = \sqrt{\frac{4\pi}{3}} \int drr^3 f_i(r) f_f(r) \int d\Omega \cdot \overline{Y_{l_f}^{m_f}} \cdot Y_1^0 \cdot Y_{l_i}^{m_i}$$

Due to the orthogonality properties of the spherical harmonics and the equation

$$Y_1^0 \cdot Y_{l_i}^{m_i} = \alpha_1 \cdot Y_{l_i+1}^{m_i} + \alpha_2 Y_{l_i-1}^{m_i}$$

(α_1, α_2 being some tabulated coefficients) we find that the matrix element is different from zero only if the conditions (selection rule)

$$l_f = l_i \pm 1, m_f = m_i$$

is fulfilled, i.e. by optical transition with linearly polarised light

$$\Delta_l = \pm 1, \Delta_m = 0$$

2. We now assume circularly polarised light propagating along z. Then

$$\vec{e}_0 = (\cos\omega t, \pm \sin\omega t, 0)$$

With the sign + (—) $\vec{e}_0$ and $+z$ build a right screw ("right circularly polarised light) respectively a left screw ("left-circularly polarized light). We compute:

$$\vec{r} \cdot \vec{e}_0 = \frac{1}{2}\left((x \pm iy) \cdot e^{-i\omega \cdot t} + (x \mp iy) \cdot e^{+i\omega \cdot t}\right)$$

From this relation we obtain the matrix elements for optical transition with circularly polarized light:

$$\sim \left(\psi_f, \hat{x} \pm i\hat{y}\psi_i\right)$$

With

$$x + iy = -r\sqrt{\frac{8\pi}{3}} Y_1^1$$

$$x - iy = r\sqrt{\frac{8\pi}{3}} Y_1^{-1}$$

we obtain the selection rules for the emission and absorption of circularly polarized light propagating in $+z$ direction:

$$\Delta_l = \pm 1$$

For right circularly polarised light

$$\Delta_m = +1$$

For left circularly polarised light

$$\Delta_m = -1$$

Example: Spin polarization of $p \to s$ optical transitions in spherically symmetric potential with spin-orbit coupling. We use e.g. left-hand polarized light propagating along $+z$, i.e. the operator is $\sim Y_1^{-1}$.

$^2S_{\frac{1}{2}}$		$Y_0^0 Y_{\frac{1}{2}}^{\frac{1}{2}}$	$Y_0^0 Y_{\frac{1}{2}}^{-\frac{1}{2}}$	
$^2P_{\frac{3}{2}}$	$Y_1^1 Y_{\frac{1}{2}}^{\frac{1}{2}}$	$\sqrt{\frac{2}{3}} \cdot Y_1^0 Y_{\frac{1}{2}}^{\frac{1}{2}} + \sqrt{\frac{1}{3}} \cdot Y_1^1 Y_{\frac{1}{2}}^{-\frac{1}{2}}$	$\sqrt{\frac{2}{3}} \cdot Y_1^0 Y_{\frac{1}{2}}^{-\frac{1}{2}} + \sqrt{\frac{1}{3}} \cdot Y_1^{-1} Y_{\frac{1}{2}}^{\frac{1}{2}}$	$Y_1^{-1} Y_{\frac{1}{2}}^{-\frac{1}{2}}$
$^2P_{\frac{1}{2}}$		$-\sqrt{\frac{1}{3}} \cdot Y_1^0 Y_{\frac{1}{2}}^{\frac{1}{2}} + \sqrt{\frac{2}{3}} \cdot Y_1^1 Y_{\frac{1}{2}}^{-\frac{1}{2}}$	$-\sqrt{\frac{1}{3}} \cdot Y_1^0 Y_{\frac{1}{2}}^{-\frac{1}{2}} + \sqrt{\frac{2}{3}} \cdot Y_1^{-1} Y_{\frac{1}{2}}^{\frac{1}{2}}$	

*The $P_{\frac{3}{2}}$ and $P_{\frac{1}{2}}$ spin-orbit spitted states build the initials states, the $S_{\frac{1}{2}}$-state build the final state. The basis states were computed using the tables for Clebsch-Gordan coefficients in M. Tanabashi et al. (Particle Data Group), Phys. Rev. D **98**, 030001 (2018)), p. 564,* `https://journals.aps.org/prd/abstract/10.1103/PhysRevD.98.030001`.

We consider for instance the $P_{\frac{3}{2}} \to S_{\frac{1}{2}}$ transition. We need to compute the 8 matrix elements

$$\langle Y_0^0 Y^{\pm}, Y_1^{-1} \cdot P_{\frac{3}{2}}^{m_j} \rangle$$

with $m_j = \frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}$. The states on the right-hand side are all possible initial states and the state $Y_{0,0} Y^{\pm}$ is the final state of the $P_{\frac{3}{2}} \to S_{\frac{1}{2}}$ transition. The light operator does not act in spin space (within the dipole approximation). To compute the matrix elements we use the simple rules

$$\hat{O}_{\text{light}} Y_l^m = c \cdot Y_{l\pm 1}^{m-1}$$
$$\hat{O}_{\text{light}} Y^{\pm} = Y^{\pm}$$

with c being some matrix element that must be computed taking the radial parts of the wave function into account. Accordingly, we have

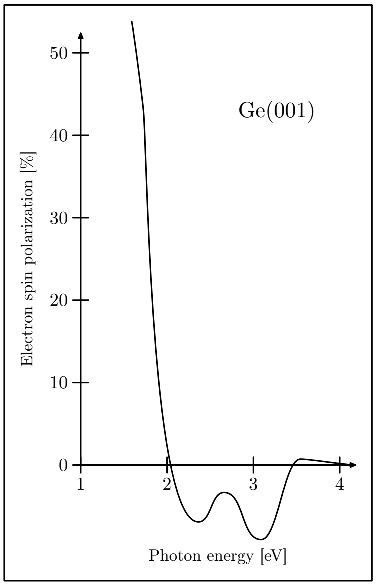
$$\langle Y_0^0, Y_1^{-1} Y_1^1 \rangle = c$$

while all others matrix elements between spherical harmonics vanish. This result implicates that the optical transition $P_{\frac{3}{2}} \to S_{\frac{1}{2}}$ involves only the two initial states $m_j = \frac{3}{2}, \frac{1}{2}$. Taking into account the weight of the various orbital wave functions within the various states (this weight being given by a Clebsch-Gordan coefficient!!) we can compute the relative strength of the transition from $m_j = \frac{3}{2}$ – leading to ↑-spins and from $m_j = \frac{1}{2}$ – leading to ↓-spins:

$$\frac{1}{\frac{1}{3}} = 3$$

This means that the optical transition $P_{\frac{3}{2}} \to S_{\frac{1}{2}}$ produces ↑ spins with a relative probability of 3 and ↓-spins with the relative probability of 1, giving rise to a spin polarization

$$\langle \sigma_z \rangle = \frac{3-1}{3+1} = 50\%$$



The spin polarization of electrons photoemitted from Ge as a function of the photon energy. At the threshold photon energy (left, where the experimental curve begins), Ge single crystals have a transition which can be described by the $P_{\frac{3}{2}} \rightarrow S_{\frac{1}{2}}$ *transition and shows the expected* 50% *spin polarization.*

11.4.2. *Periodic perturbation with finite frequency width (technical).

Despite laser light being available almost as monochromatic source, one obtains more physical results if one assigns a finite frequency width to a given source of light. In this case,

$$\vec{E}(t) = \frac{1}{2\pi}\int d\omega\left(\vec{\tilde{E}}(\omega)e^{-i\omega t} + \overline{\vec{\tilde{E}}(\omega)}e^{+i\omega t}\right)$$

The Fourier component is, in general, a linear combination of right and left circularly polarized. We study for instance the excitation from E_1 to $E_2 > E_1$with right circularly polarized light:

$$\vec{\tilde{E}}(\omega) = E_+(\omega)\cdot(\vec{e}_x + i\vec{e}_y)$$

The absorption process is governed by the $e^{-i\omega t}$ components of the field. The matrix element writes then

$$\frac{1}{2\pi}\int d\omega\big(\psi_2, -E_+(\omega)\cdot(\vec{e}_x + i\vec{e}_y)e^{-i\omega t}\cdot\vec{D}\psi_1\big) = -\big(\psi_2, \vec{D}_x + i\vec{D}_y\psi_1\big)\cdot\frac{1}{2\pi}\int d\omega E_+(\omega)e^{-i\omega t}$$

As a consequence of the light beam being treated as a superposition of plane waves to form a narrowly shaped beam in the frequency regime, we do not have any mathematical problems in letting $T \to \infty$, i.e. in letting the perturbation act for an infinite amount of time. The transition probability writes

$$|\, C_{1\to 2}\,|^2 = \frac{1}{\hbar^2}\cdot\underbrace{|\int_{-\infty}^{\infty} dt\cdot\frac{1}{2\pi}\int d\omega E_+(\omega)e^{-i\omega t}\cdot e^{\frac{i}{\hbar}(E_2 - E_1)\cdot t}\,|^2}_{|E_+(E_2 - E_1)|^2}\,|\big(\psi_2, (\vec{D}_x + i\vec{D}_y)\psi_1\big)\,|^2$$

$$= \frac{1}{\hbar^2}\,|\, E_+(\frac{E_2 - E_1}{\hbar})\,|^2|\big(\psi_2, (\vec{D}_x + i\vec{D}_y)\psi_1\big)\,|^2$$

The first factor is the Fourier transform of the light beam taken at a frequency which corresponds to the energy difference $E_2 - E_1$, in line with the Bohr frequency condition. The transition from the level E_2 to E_1 with light emission is driven by the $e^{i\omega t}$ components has the probability

$$|\, C_{1\to 2}\,|^2 = \frac{1}{\hbar^2}\,|\,\overline{E_+(\frac{E_2 - E_1}{\hbar})}\,|^2|\,\overline{\big(\psi_2, (\vec{D}_x + i\vec{D}_y)\psi_1\big)}\,|^2$$

in line with the identity of Einstein coefficients. The pre-factor

$$|\, E_+(\frac{E_2 - E_1}{\hbar})\,|^2$$

is related to the intensity of the light at the frequency $\frac{E_2 - E_1}{\hbar}$. More precisely: The total energy transported by the light wave per unit time interval and unit area is

$$J = \epsilon_0\cdot c\frac{1}{T}\int_{-\infty}^{\infty} dt\vec{E}(t)\cdot\overline{\vec{E}(t)}$$

Using

$$\vec{E}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\vec{E}(\omega)e^{-i\omega t}$$

and inserting this expression we obtain

$$J = \epsilon_0\cdot c\frac{1}{T}\int_{-\infty}^{\infty} dt\frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\vec{E}(\omega)e^{-i\omega t}\cdot\overline{\frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega'\vec{E}(\omega')e^{+i\omega' t}}$$

$$= \epsilon_0\cdot c\cdot\frac{1}{T}\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\vec{E}(\omega)\overline{\vec{E}(\omega)}$$

$$= \epsilon_0\cdot c\cdot\frac{1}{T}\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\cdot\frac{1}{2}\cdot\left(E_+(\omega)\frac{1}{\sqrt{2}}(\vec{e}_x + i\vec{e}_y) + \overline{E_+(\omega)\frac{1}{\sqrt{2}}(\vec{e}_x + i\vec{e}_y)}\right)^2$$

$$= \epsilon_0\cdot c\cdot\frac{1}{T}\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\,|\, E_+(\omega)\,|^2 = \epsilon_0\cdot c\cdot\frac{1}{T}\cdot\frac{1}{\pi}\int_{0}^{\infty} d\omega\,|\, E_+(\omega)\,|^2$$

$|E_+(\omega)|^2$ is related to the total energy transported per unit time, per unit area and per unit frequency $\frac{dJ}{d\omega}$:

$$|E_+(\omega)|^2 = 4\pi^2 \cdot \frac{\frac{dJ}{d\omega} \cdot T}{4\pi\epsilon_0 c}$$

and obtain

$$\Gamma_{1\to 2} = 4\pi^2 \cdot \frac{\frac{dJ}{d\omega}\left(\frac{E_2 - E_1}{\hbar}\right)}{4\pi\epsilon_0 c\hbar^2} \left|\left(\psi_2, (\vec{D}_x + i\vec{D}_y)\psi_1\right)\right|^2$$

This expression avoids using the Dirac delta function in the expression for the transition rate. The relevant component of the transition rate – the matrix element – is maintained also in this physically more accurate expression.

Part V.

Exercises

12. Exercises.

12.1. Exercises to Chapter 1.

In this section, we solve some simple exercises that allow us to become acquainted with quantities and units used in quantum mechanics.

1. De Broglie Wavelength.

According to de Broglie the wavelength of a free particle is given by

$$\lambda = \frac{2\pi\hbar}{\sqrt{2m\cdot E_{kin}}}$$

E_{kin} being the kinetic energy of the particle and m the mass. Use this relation to express the wavelength of an electron as

$$\lambda[nm] = \sqrt{\frac{1.504}{E_{kin}[eV]}}$$

In this formula, $E_{kin}[eV]$ and $\lambda[nm]$ means that if you plug in a number representing E_{kin} in eV you obtain a number representing λ in nm.

Solution.

Insert

$$\hbar = 6.582\cdot 10^{-16} \quad \text{eV}\cdot\text{sec}$$

and

$$m = 0.911\cdot 6.242\cdot 10^{-30}\cdot\frac{(eV)^2\cdot sec^2}{(nm)^2}$$

(the unit of kg for the mass has been converted in suitable units) into De-Broglie relation to obtain the desired useful formula.

2. Energy of the ground state of Hydrogen.

The Bohr model foresees that the energy of the ground state of the hydrogen is given by

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

with $n = 1$. Compute this energy in eV.

Solution

Use

$$e \approx 1.6\cdot 10^{-19} A\cdot sec \quad m \approx 9.11\cdot 10^{-31} kg$$

$$\epsilon_0 \approx 8.85\cdot 10^{-12}\frac{A\cdot sec}{V\cdot m} \quad \hbar \approx 1.05\cdot 10^{-34} J\cdot sec$$

Inserting these values one obtains

$$E_1 \approx -2.2 \cdot 10^{-18} \text{Joule}$$

as

$$1 \quad \text{Joule} = 6.24210^{18} eV$$

one obtains

$$E_1 \approx -13.6 eV$$

3. Classical circular orbits.

Bohr computation of the quantized energy levels of the Hydrogen atom uses results from the Newton mechanics of circular orbits in the radially symmetric Coulomb potential

$$-\frac{1}{4\pi\epsilon_0} \cdot \frac{e^2}{r}$$

(r being the radial coordinate).

- 1. Find the radius of the circular orbits by minimizing the effective radial potential energy

$$U_{eff}(r) = -\frac{1}{4\pi\epsilon_0} \cdot \frac{e^2}{r} + \frac{L^2}{2mr^2}$$

- 2. Insert the radius in the expression for the total energy to obtain the relation between E and L used by Bohr.

Solution

- 1.

$$\frac{d}{dr} U_{eff}(r) = \frac{1}{4\pi\epsilon_0} \cdot \frac{e^2}{r^2} - \frac{L^2}{mr^3} =!0 \Leftrightarrow r = \frac{L^2 \cdot 4\pi\epsilon_0}{m \cdot e^2}$$

- 2. Inserting

$$r = \frac{L^2 \cdot 4\pi\epsilon_0}{m \cdot e^2}$$

in

$$E = -\frac{1}{4\pi\epsilon_0} \cdot \frac{e^2}{r} + \frac{L^2}{2mr^2}$$

produces

$$E = -\frac{1}{4\pi\epsilon_0} \cdot \frac{e^2}{r} + \frac{L^2}{2mr^2} = -\frac{1}{4\pi\epsilon_0} \cdot \frac{e^2 m e^2}{L^2 4\pi\epsilon_0} + \frac{L^2 m^2 e^4}{2m L^4 (4\pi)^2 \epsilon_0^2} =$$
$$-\frac{1}{(4\pi\epsilon_0)^2} \cdot \frac{e^4 m}{L^2} + \frac{m e^4}{2L^2 (4\pi\epsilon_0)^2} = -\frac{1}{2} \cdot \frac{e^4}{(4\pi\epsilon_0)^2} \cdot \frac{m}{L^2}.$$

4. Wavelength of a Photon.

Compute the wavelength λ of a photon in nm as a function of the photon energy E given in eV.

Solution.

From the quantum relation

$$E = \hbar \cdot \omega$$

and the classical relation for waves

$$\omega = \frac{2\pi}{\lambda} \cdot c$$

(c: light velocity) one finds

$$\lambda = \frac{\hbar 2\pi c}{E}$$

Inserting

$$\hbar = 6.582 \cdot 10^{-16} \quad \mathrm{eV} \cdot \mathrm{sec} \quad \text{and} \quad c = 3 \cdot 10^{17} \frac{nm}{sec}$$

one finds the ”easy to remember” relation

$$\lambda[nm] = \frac{1240}{E[eV]}$$

12.2. Exercises to Chapters 2,3,4.

1. Schrödinger current.

Use the time independent Schrödinger equation for the motion of a mass m along the x-axis in a general potential $\Phi(x)$ to show that

$$\frac{d}{dx}\left[\frac{\hbar}{2im}\Big(\bar{\psi}(x)\frac{d}{dx}\psi(x) - \psi(x)\frac{d}{dx}\bar{\psi}(x)\Big)\right] = 0$$

Solution.

Solution:

$$\begin{aligned}
\frac{d}{dx}\left[\frac{\hbar}{2im}\Big(\bar{\psi}(x)\frac{d}{dx}\psi(x) - \psi(x)\frac{d}{dx}\bar{\psi}(x)\Big)\right] &= \frac{\hbar}{2im} \cdot \left(\bar{\psi}(x)\frac{d^2}{dx^2}\psi(x) - \psi(x)\frac{d^2}{dx^2}\bar{\psi}(x)\right) \\
&= \frac{\hbar}{2im} \cdot \frac{2m}{\hbar^2} \cdot \Big(-(E - \Phi(x)) \cdot \overline{\psi(x)}\psi(x) + (E - \Phi(x)) \cdot \psi(x)\overline{\psi(x)}\Big) \\
&= 0
\end{aligned}$$

2. Matrix algebra.

This is an exercise to remind of some notions of matrix algebra. Given is the matrix

$$R := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

$\varphi \in [0, 2\pi]$. This matrix represents rotations in the xy-plane around the z-axis.

1. Show that this matrix is not hermitic.
2. Compute the eigenvalues of this matrix.
3. Compute the eigenvectors.
4. Compute the transformation of the vectors $(1,0)$ and $(0,1)$ under the action of this matrix.
5. As a generalization of this exercise, consider a rotation of any vector in our three dimensional space by an angle φ around a given axis specified by the unit vector $\mathbf{n}$. Do not try to find the general matrix representation of this rotation. Instead, define one, very particular, coordinate system that allows to express the matrix representation of this rotation with as many vanishing matrix elements as possible and write the matrix representation in this coordinate system. Compute the eigenvalues of this matrix and find the eigenvectors (use, if pertinent, the results of the previous questions).

Solution.

1.
$$\overline{R^T} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \neq R$$

2. To find the eigenvalues of R we have to solve
$$\begin{aligned} \det\begin{pmatrix} \cos\varphi-\lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi-\lambda \end{pmatrix} &= (\cos\varphi-\lambda)(\cos\varphi-\lambda)-(-\sin\varphi\sin\varphi) \\ &= \cos^2\varphi-2\lambda\cos\varphi+\lambda^2+\sin^2\varphi \\ &= \lambda^2-2\lambda\cos\varphi+1 \\ != & \;0 \end{aligned}$$
The Eigenvalues of R are the solutions of the equation
$$\lambda^2-2\lambda\cos\varphi+1=0$$
i.e.
$$\lambda_{1,2}=e^{\pm i\varphi}$$

3. The eigenvectors to $e^{i\varphi}$ are the solution of
$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = e^{i\varphi}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$
The eigenspace is one-dimensional, i.e. we can set $v_1=1$ and solve e.g. the equation
$$\cos\varphi\cdot 1-\sin\varphi v_2=e^{i\varphi}\cdot 1$$
to find
$$v_2=-i$$
Accordingly, the eigenspace to $e^{i\varphi}$ consits of all vectors parallel to
$$\begin{pmatrix} 1 \\ -i \end{pmatrix}.$$
In a similar manner we find that the eigenspace to $e^{-i\varphi}$ consists of all vectors parallel to
$$\begin{pmatrix} 1 \\ i \end{pmatrix}$$

4.
$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$$
$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}$$
i.e. the columns of the matrix contain the transformed vectors.

5. For solving this task we need to find a suitable coordinate system. If one choose the z axis so that $\mathbf{n}$ is along z, we can use the results from the previous matrix and write
$$R(\varphi)=\begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
For calculating the eigenvalues we need to solve the determinantal equation
$$\det(R-\lambda\mathbb{1})=0$$
We observe that the rotation matrix is in block form, so that the determinant can be computed easily and the determinantal equation writes

$$(\lambda^2-2\lambda\cos\varphi+1)\cdot(1-\lambda)=0$$

The solutions are the three eigenvalues

$$e^{i\varphi} \quad e^{-i\varphi} \quad 1$$

and the eigenvectors

$$\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}.$$

$$\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}.$$

and

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

respectively.

3. The Dirac delta function, seen by W. Pauli (Pauli Lectures on Physics, Vol. 5: Wave Mechanics).

Consider the set of functions defined by the following conditions:

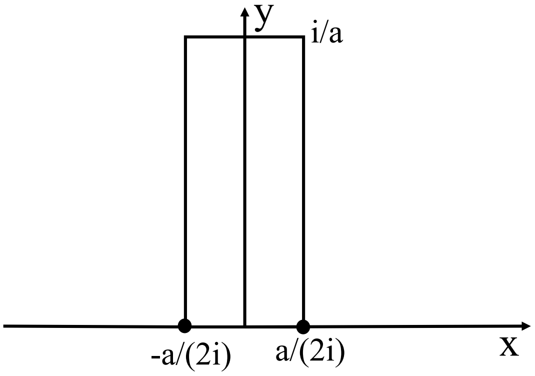
$$\delta_i(x)=\begin{cases} 0 & |x|>\frac{a}{i\cdot 2} \\ \frac{i}{a} & |x|<\frac{a}{i\cdot 2} \end{cases}$$

$i=1,2,\ldots.$

1. Make yourself familiar with these functions by plotting their graph for some values of i.
2. Show that
$$\int_{-\infty}^{\infty}\delta_i(x)dx=1$$
for any value of i.
3. Compute
$$\lim_{i\to\infty}\delta_i(x)$$
for $x=0$ and $x\neq 0$.
4. Now the difficult part. Consider any function $f(x)$ which is continuous in $x=0$. Show that
$$\lim_{i\to\infty}\int_{-\infty}^{\infty}\delta_i(x)\cdot f(x)\cdot dx=f(0)$$
Hint: make use of the following exact result of Analysis (which probably you do not know, but now you learn it):
$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_{-\frac{\epsilon}{2}}^{\frac{\epsilon}{2}}f(x)dx=f(0)$$
If you have some time, you can even demonstrate this result as an application of your knowledge of Analysis.

Solution.

1. The graph of one such functions is given in the figure.



The graph of a rectangular function. The graph of the Dirac-delta functions can be obtained by letting the index i go to infinity.

2. The area of the rectangles is
$$\frac{i}{a}\cdot\frac{a}{i}=1 \quad \text{independent of} \quad i$$
3. $\delta_i(x)$ for $x=0$ tends to infinity for $i\to\infty$. $\delta_i(x)$ tends to 0 for $x\neq 0$ for $i\to\infty$.
4. $$\int_{-\infty}^{\infty}\delta_i(x)\cdot f(x)\cdot dx \;=\; \frac{i}{a}\cdot\int_{-\frac{a}{2i}}^{\frac{a}{2i}}dx\cdot f(x)=\frac{1}{\frac{a}{i}}\cdot\int_{-\frac{a}{2i}}^{\frac{a}{2i}}dx\cdot f(x)$$
   Now you can use the analysis result to obtain (set $\frac{a}{i}=\epsilon$)
   $$\lim_{i\to\infty}\frac{1}{\frac{a}{i}}\cdot\int_{-\frac{a}{2i}}^{\frac{a}{2i}}dxf(x)=f(0)$$
   A final comment: when you have shown all these points, than you have learned that
   $$\lim_{i\to\infty}\delta_i(x)=\delta(x)$$
   i.e. you have found a further representation of the Dirac delta functions. You can imagine: there are many representations of the Dirac delta function, in fact there is an infinite number of them, because "A distribution is a class of equivalent testseries $\{t_n(x)\}$" (this is how Distributions are often defined in mathematics).

4. f-level of the hydrogen Atom.

The f-level of an hydrogen atom contains the 7 wave functions
$$f(r)\cdot Y_3^3(\theta,\phi)\quad f(r)\cdot Y_3^2(\theta,\phi)\quad f(r)\cdot Y_3^1(\theta,\phi)$$
$$f(r)\cdot Y_3^0(\theta,\phi)\quad f(r)\cdot Y_3^{-1}(\theta,\phi)\quad f(r)\cdot Y_3^{-2}(\theta,\phi)$$
$$f(r)\cdot Y_3^{-3}(\theta,\phi)$$
$f(r)$ is some radial dependent, normalized function and $Y(\theta,\phi)$ are spherical harmonics. Write the matrix representation of L_z and $\mathbf{L}^2$ in the f-subspace.

Solution.

We must compute the matrix elements

$$< Y_3^m, L_z Y_3^m > \quad < Y_3^m, L^2 Y_3^m > \quad m = 3, 2, 1, 0, -1, -2, -3$$

We find

$$L_z = \begin{pmatrix} \hbar\cdot 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \hbar\cdot 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \hbar\cdot 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \hbar\cdot 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \hbar\cdot(-1) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \hbar\cdot(-2) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \hbar\cdot(-3) \end{pmatrix}$$

$$\mathbf{L}^2 = \begin{pmatrix} \hbar^2\cdot 12 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \hbar^2\cdot 12 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \hbar^2\cdot 12 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \hbar^2\cdot 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \hbar^2\cdot 12 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \hbar^2\cdot 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \hbar^2\cdot 12 \end{pmatrix}$$

5. Inversion operator and Hamilton operator for a mass in an infinite potential well.

A particle undergoes a motion in the interval $[-\frac{L}{2}, \frac{L}{2}]$ along the x-axis. The motion is determined by the Schrödinger equation. The boundary conditions are the vanishing of the wave function at the edges of the interval. The potential energy is taken to be zero inside the interval, infinite outside. The inversion operator is defined by

$$\mathscr{I}\psi(x) = \psi(-x)$$

1. Show that $\mathscr{I}$ is Hermitic.
2. Find the eigenvalues of $\mathscr{I}$.
3. Find a CONS of eigenfunctions of $\mathscr{I}$.
4. Which energy-eigenvalue and which eigenfunction describe the ground state of the system?

Assume now that an experimental device has localized the particle within the interval in a state

$$\psi(x) = \frac{1}{\sqrt{L}}$$

5. Which values are the possible results of measuring the observable related to the operator $\mathscr{I}$ in this state?
6. What is the probability of measuring these values?
7. Compute the expectation value of $\mathscr{I}$ in this state.
8. Which values are the possible results of measuring the energy in this state?
9. What is the probability of measuring these values?
10. Compute the expectation value of the energy in this state.

Solution.

1. $$\begin{aligned} (\mathscr{I}\psi,\psi) &= \int_{-\infty}^{\infty} \overline{\psi(-x)}\cdot\psi(x)\,dx \underbrace{=}_{y=-x,dx=-dy} -\int_{\infty}^{-\infty} \overline{\psi(y)}\psi(-y)\,dy \\ &= \int_{-\infty}^{\infty} \overline{\psi(y)}\mathscr{I}\psi(y)\,dy \\ &= (\psi,\mathscr{I}\psi) \end{aligned}$$

2. We use
$$\mathscr{I}^2 = \mathbb{1}$$
to transform the eigenvalue equation
$$\mathscr{I}\psi(x) = \lambda\cdot\psi(x)$$
Into
$$(\mathscr{I})^2\psi(x) = \psi(x) = (\lambda)^2\psi(x)$$
so that
$$\lambda^2 = 1 \quad \lambda = \pm 1$$

3. The elements of the CONS have to fullfill two conditions:
First: They have to fullfill the SE for the given potential and boundary conditions.
Second: They have to be eigenfunctions of the specific operator we are looking at.
For the **first** part we see that the potential between $[-\frac{L}{2},\frac{L}{2}]$ is vanishing and that the wavefunction has to be zero at the edges. This reminds us of the infinite 1d potential well, so that we can use the results from Chapter 2 and find the set of basis functions
$$p_n(x) = \sqrt{\frac{2}{L}}cos\left(\frac{(2n+1)\pi}{L}x\right) \quad n = 0,1,2,...$$
and
$$q_m(x) = \sqrt{\frac{2}{L}}sin\left(\frac{(2m)\pi}{L}x\right) \quad m = 1,2,3,...$$
For the **second** part we apply $\mathscr{I}$ to p_n and q_m in order to check if they are eigenfunctions or not. Indeed
$$\mathscr{I}p_n(x) = 1\cdot p_n(x)$$
and
$$\mathscr{I}q_m(x) - 1\cdot q_m(x).$$

4. $$E_0 = \frac{\hbar^2}{2m}\cdot\frac{\pi^2}{L^2} \quad p_0(x) = \sqrt{\frac{2}{L}}cos\left(\frac{\pi}{L}x\right)$$

5. The wave function is a constant in the interval. A constant is even with respect to change of sign and therefore it is an eigenstate of $\mathscr{I}$ to the eigenvalue $+1$.

6. The probability of measuring $(+1)$ is 1.

7. The expectation value is
$$<\mathscr{I}>|_{\psi} = 1\cdot(+1) = +1$$

8. The constant function is not an eigenstate of the Hamilton operator and does not have a trivial expansion into the eigenstates p_n and q_m. Accordingly, in principle, the possible outcomes of measuring the energy are all eigenvalues, with some probability.

9. The amplitude of measuring the eigenvalues E_n amounts to

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} dx\cdot\frac{1}{\sqrt{L}}\cdot\sqrt{\frac{2}{L}}\cos(\frac{(2n+1)\pi}{L}x)=\frac{2\sqrt{2}\cdot(-1)^n}{\pi\cdot(2n+1)}$$

The amplitude of measuring the eigenvalues E_m amounts to

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} dx\cdot\frac{1}{\sqrt{L}}\cdot\sqrt{\frac{2}{L}}\sin(\frac{(2m)\pi}{L}x)=0$$

$n=0,1,2,...$ and $m=1,2,......$ Both amplitudes were computed using Wolfram Alpha software, available online.

10. The expectation value of E amounts the

$$\frac{\pi^2\hbar^2}{2mL^2}\sum_{n=0}^{\infty}(2n+1)^2\cdot\frac{8}{\pi^2\cdot(2n+1)^2}=\infty$$

This anomalous result is due to the fact that there is an infinite number of energy bound states which can be assumed with a high enough probability.

6. Particle in a Quadratic Potential.

A point mass m moves according to the laws of quantum mechanics in space $\mathscr{R}^3$ and is subject to the potential energy

$$\Phi(x,y,z)=\frac{1}{2}\cdot\alpha\cdot\left(x^2+y^2+z^2\right)$$

$\alpha>0$ being a "spring" constant.

1. Write the time independent Schrödinger equation of the particle, using cartesian coordinates.
2. Use the separation Ansatz
$$\psi(x,y,z)=f(x)\cdot g(y)\cdot h(z)$$
to write the SG as the sum of four terms, the first depending only from x, the second only from y, the third depending only from z and the 4th being exactly the sought for energy eigenvalue E. Set then term $''1''$ equal to the separation constant λ_x, term $''2''$ equal to the separation constant λ_y and term $''3''$ equal to the separation constant λ_z.
3. You obtain in this way separate ordinary differential equations in x, y and z. Find the separation constants by solving these differential equations. HINT: You do not have to solve these differential equations explicitly, you have already seen the solution in the lecture.
4. Find all possible energy eigenvalues of the original problem.
5. How many quantum numbers are necessary to describe the energy eigenvalues?

Solution.

1.
$$\frac{-\hbar^2}{2m}\nabla^2\psi(x,y,z)+\frac{\alpha}{2}\left(x^2+y^2+z^2\right)\psi-E\cdot\psi(x,y,z)=0$$

2. We plug the given Ansatz into the Schrödinger equation from a):
$$\frac{-\hbar^2}{2m}\frac{\partial^2 f(x)}{\partial x^2}g(y)\cdot h(z)-\frac{\partial^2 g(y)}{\partial y^2}f(x)\cdot h(z)-\frac{\partial^2 h(z)}{\partial z^2}f(x)\cdot g(y)+$$

$$\frac{\alpha}{2}\left(x^2+y^2+z^2\right)f(x)\cdot g(y)\cdot h(z)-E\cdot f(x)\cdot g(y)\cdot h(z)=0$$

By dividing with $\psi(x,y,z)=f(x)\cdot g(y)\cdot h(z)$ and rearranging the terms we obtain:

$$\underbrace{\frac{\hbar^2}{2m}\frac{\frac{d^2f}{dx^2}}{f}-\frac{\alpha}{2}x^2}_{=\lambda_x}+\underbrace{\frac{\hbar^2}{2m}\frac{\frac{d^2g}{dy^2}}{g}-\frac{\alpha}{2}y^2}_{=\lambda_y}+\underbrace{\frac{\hbar^2}{2m}\frac{\frac{d^2h}{dz^2}}{h}-\frac{\alpha}{2}z^2}_{=\lambda_z}+E=0.$$

This equation must hold for any x,y,z. It can only do it if each summand is equal to a constant: the separation constant.

3. By the separation Ansatz, the original SE is divided into four separate equations:

$$-\frac{\hbar^2}{2m}\frac{d^2f}{dx^2}+\frac{\alpha}{2}x^2\cdot f=-\lambda_x f$$
$$-\frac{\hbar^2}{2m}\frac{d^2g}{dy^2}+\frac{\alpha}{2}y^2\cdot g=-\lambda_y g$$
$$-\frac{\hbar^2}{2m}\frac{d^2h}{dz^2}+\frac{\alpha}{2}z^2\cdot g=-\lambda_z g$$
$$E=-\lambda_x-\lambda_y-\lambda_z$$

From the lecture, we know that the three ordinary differential equations describe quantum mechanical harmonic oscillators, so that we know their solutions:

$$\Rightarrow\quad \lambda_x=-\hbar\sqrt{\frac{\alpha}{m}}\left(n_x+\frac{1}{2}\right)\quad \lambda_y=-\hbar\sqrt{\frac{\alpha}{m}}\left(n_y+\frac{1}{2}\right)\quad \lambda_z=-\hbar\sqrt{\frac{\alpha}{m}}\left(n_z+\frac{1}{2}\right)$$

with

$$n_x=0,1,2,\ldots\quad n_y=0,1,2,\ldots\quad n_z=0,1,2,\ldots.$$

4. Inserting the separation constants we find

$$\begin{aligned}E(n_x,n_y,n_z) &= -\lambda_x-\lambda_y-\lambda_z=\hbar\sqrt{\frac{\alpha}{m}}\left(n_x+\frac{1}{2}\right)+\hbar\sqrt{\frac{\alpha}{m}}\left(n_y+\frac{1}{2}\right)+\hbar\sqrt{\frac{\alpha}{m}}\left(n_z+\frac{1}{2}\right)=\\ &= \hbar\sqrt{\frac{\alpha}{m}}\left(n_x+n_y+n_z+\frac{3}{2}\right)\end{aligned}$$

5. To describe the eigenvalues of the energy we need the 3 quantum numbers n_x,n_y,n_z.

7. Particle in the xy-plane with infinite potential walls.

A particle with mass m moves according to the Schrödinger equation in the xy-plane. Within a the square region

$$-\frac{L}{2}\leq x\leq\frac{L}{2}\text{and}-\frac{L}{2}\leq y\leq\frac{L}{2}$$

the potential is zero. Outside this region the potential is infinite.

•. Write the time independent Schrödinger equation for the particle in the square region.

• Determine the boundary conditions.

• Find the eigenvalues.

Solution.

-
$$\left[\frac{-\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\right]\psi(x,y)=E\cdot\psi(x,y)$$

- $$\psi\left(x=\pm\frac{L}{2},y\right)=0 \quad \text{and} \quad \psi\left(x,y=\pm\frac{L}{2}\right)=0 \quad .$$

- Set

$$\psi(x,y)=f(x)\cdot g(y)$$

and insert in the SE:

$$\begin{aligned}\left[\frac{-\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\right]f(x)g(y) &= Ef(x)g(y) \quad ||:f(x)\cdot g(y) \\ \underbrace{\frac{\hbar^2}{2m}\frac{f''}{f}}_{=-\lambda_x}+\underbrace{\frac{\hbar^2}{2m}\frac{g^{\cdot\cdot}}{g}}_{=-\lambda_y}+E &= 0 \quad .\end{aligned}$$

The separation Ansatz has divided the original SE into a differential equation for the separation constant λ_x, a differential equation for the separation constant λ_y and an algebraic equation

$$E=\lambda_x+\lambda_y$$

λ_x and λ_y are the eigenvalues of differential equations for the one one-dimensional infinite walls in the variables x respectively y and are therefore known to be

$$\lambda_x=\frac{\pi^2\hbar^2}{2mL^2}(n_x)^2 \quad n_x=1,2,...$$

and

$$\lambda_y=\frac{\pi^2\hbar^2}{2mL^2}(n_y)^2 \quad n_y=1,2,...$$

The algebraic equation leads to

$$E(n_x,n_y)=\frac{\pi^2\hbar^2}{2mL^2}[n_x^2+n_y^2] \quad n_x=1,2,... \quad n_y=1,2,...$$

8. Separation Ansatz.

Consider the time-independent Schrödinger equation in three dimensions for the wave function $\psi(x,y,z)$. Let the potential energy depend only on the variable x,y. Along z assume periodic boundary conditions in the interval $[-\frac{L}{2},\frac{L}{2}]$. Use your knowledge from the lectures to write the most accurate separation Ansatz.

Solution.

As the potential is only dependent on x,y, a possible separation Ansatz reads

$$\psi(x,y,z)=f(x,y)\cdot g(z)$$

i.e. the z-component can be separated out. As along z we have periodic boundary conditions, we can specify our Ansatz to

$$\psi(x,y,z)=f(x,y)\cdot g(z)\cdot\frac{1}{\sqrt{L}}\cdot e^{ik_n\cdot z} \quad k_n=0,\pm\frac{2\pi}{L},\pm 2\cdot\frac{2\pi}{L},.....$$

9. Particle in an infinite potential wall.

In the lecture we have solved the problem of a particle with mass m subject to infinite potential walls:

$$\Phi(x)=\begin{cases}\infty; & L/2\leq x\leq -L/2\\ 0; & -L/2\leq x\leq L/2\end{cases}$$

1. Find in the manuscript the wave function describing the ground state of the particle

2. What is the amplitude of finding the particle at $x=0$ in the ground state?
3. What is the probability density of finding the particle at $x=0$ in the ground state?
4. Compute the probability of finding the particle in an interval $[-\epsilon,\epsilon]$ in the ground state
5. Find in the manuscript the wave function describing the lowest lying excited state
6. What is the amplitude of finding the particle at $x=0$ in this state?
7. What is the probability density of finding the particle at $x=0$ in this state?
8. Compute the probability of finding the particle in an interval $[-\epsilon,\epsilon]$ in this state

Solution.

We recall from the lecture the general solution of the problem:

$$\sqrt{\frac{2}{L}}\cos\left(\frac{(2n+1)\pi}{L}x\right) \quad E_n=\frac{\pi^2\hbar^2}{2mL^2}(2n+1)^2 \quad n=0,1,2,...$$

and

$$\sqrt{\frac{2}{L}}\sin\left(\frac{(2n)\pi}{L}x\right) \quad E_p=\frac{\pi^2\hbar^2}{2mL^2}(2p)^2 \quad p=1,2,3,...$$

1. The ground state is the solution with the lowest possible energy, i.e.

$$n=0 \quad E_0=\frac{\pi^2\hbar^2}{2mL^2} \quad \psi_0(x)=\sqrt{\frac{2}{L}}\cos(\frac{\pi}{L}x)$$

2.

$$\psi_0(x_0=0)=\sqrt{\frac{2}{L}}$$

3.

$$|\psi_0(x_0)|^2=|\sqrt{\frac{2}{L}}|^2=\frac{2}{L}$$

4.

$$W_{\psi_0}(x\in[-\epsilon,\epsilon])=\int_{-\epsilon}^{\epsilon}\left(\sqrt{\frac{2}{L}}\cdot\cos(\frac{\pi}{L}x)\right)^2\cdot dx=\frac{2}{L}\epsilon+\frac{sin\left(\frac{2\pi\epsilon}{L}\right)}{\pi}$$

5. The lowest lying excited state is

$$p=1 \quad E_{p=1}=\frac{\pi^2\hbar^2}{2mL^2}(2)^2. \quad \psi_1(x)=\sqrt{\frac{2}{L}}\sin(\frac{2\pi}{L}x)$$

6. The amplitude is $\psi_1(x_0=0)=\sqrt{\frac{2}{L}}\sin(\frac{2\pi}{L}0)=0$.
7. The probability density in the point $x_0=0$ is given by: $\rho=|\psi_1(x_0)|^2=|\sqrt{\frac{2}{L}}\sin(\frac{2\pi}{L}0)|^2=0$.
8.

$$W_{\psi_1}(x\in[-\epsilon,\epsilon])=\frac{\int_{-\epsilon}^{\epsilon}|\psi_1(x)|^2\,dx}{\int_{-L/2}^{L/2}|\psi_1(x)|^2\,dx}=\frac{2}{L}\epsilon-\frac{sin\left(\frac{4\pi\epsilon}{L}\right)}{2\pi}$$

10. Free particle with periodic boundary conditions.

The Fourier theorem states that the superposition of plane waves

$$u_n(x) = \frac{1}{\sqrt{L}} \cdot e^{i\frac{2\pi\cdot n}{L}\cdot x} \quad n = 0, \pm 1, \pm 2, ...$$

can be used to express any L periodic function in the interval $[\frac{-L}{2}, \frac{L}{2}]$

- Show that the plane waves are eigenfunctions of the momentum operator and find all possible eigenvalues of the momentum operator.

Assume now that one has prepared, by a suitable experimental device, a particle in a state with wave function

$$\psi(x) = \frac{1}{\sqrt{3}} \frac{1}{\sqrt{L}} e^{i\frac{2\pi}{L}\cdot x} + \frac{i\cdot\sqrt{2}}{\sqrt{3}} \frac{1}{\sqrt{L}} e^{-i\frac{2\pi}{L}\cdot x}$$

- Find all possible outcomes of an experiment that measures the momentum in the state $\psi(x)$.
- With which probability are these values detected? Hint: you do not have to compute one single integral explicitely for solving this exercise! Use the orthonormality relations to compute any scalar product forthcoming in the computation.
- Compute the quantum mechanical expectation value of the momentum operator in the state ψ by computing the scalar product $\left(\psi, p_{op}\psi\right)$.
- Use now the possible outcomes of the measurement of the momentum and their probabilities to compute again the quantum mechanical expectation value.
- Find all possible outcomes of an experiment that measures the energy in the state $\psi(x)$.
- Compute the probability of these outcomes.
- Find the expectation value of the energy operator in this state.

Solution.

-

$$-i\hbar\frac{d}{dx}\frac{1}{\sqrt{L}} \cdot e^{i\frac{2\pi\cdot n}{L}\cdot x} = \underbrace{\hbar\cdot\frac{2\pi\cdot n}{L}}_{\text{eigenvalue}} \cdot \underbrace{\frac{1}{\sqrt{L}} \cdot e^{i\frac{2\pi\cdot n}{L}\cdot x}}_{\text{eigenfunction}}$$

- The given function $\psi(x)$ is a linear combination of two eigenfunctions of the momentum operator: $u_1(x)$ and $u_{-1}(x)$. From the lecture, we know that the detection of the momentum of the particle in this state will produce the two corresponding eigenvalues:

$$\pm\hbar\frac{2\pi}{L}$$

- The given function is normalized to "1" so that

$$\begin{aligned} W_\psi(\hbar\frac{2\pi}{L}) &= |<u_1, \psi>|^2 = |<u_1, \frac{1}{\sqrt{3}}u_1 + \frac{i\sqrt{2}}{\sqrt{3}}u_{-1}>|^2 \\ &= |<u_1, \frac{1}{\sqrt{3}}u_1> + <u_1, \frac{i\sqrt{2}}{\sqrt{3}}u_{-1}>|^2 = |\frac{1}{\sqrt{3}}<u_1, u_1> + \frac{i\sqrt{2}}{\sqrt{3}}<u_1, u_{-1}>|^2 \\ &= |\frac{1}{\sqrt{3}}<u_1, u_1> + 0|^2 = |\frac{1}{\sqrt{3}}|^2 = \frac{1}{3} \end{aligned}$$

and

$$W_\psi(-\hbar\frac{2\pi}{L}) = |<u_{-1}, \psi>|^2 = |\frac{i\sqrt{2}}{\sqrt{3}}<u_{-1}, u_{-1}>|^2 = \frac{2}{3}$$

- The wave function is normalized to "1". Using the orthonormality properties of the plane waves we find

$$\hat{p}\psi(x) = -i\hbar\frac{d}{dx}\psi(x) = \hbar\frac{2\pi}{L}\cdot\frac{1}{\sqrt{3}}u_1 - \hbar\frac{2\pi}{L}\cdot\frac{i\cdot\sqrt{2}}{\sqrt{3}}u_{-1}$$

and

$$\begin{aligned}(\psi,\hat{p}\,\psi) &= \hbar\frac{2\pi}{L}\cdot\left(\frac{1}{\sqrt{3}}u_1+\frac{i\cdot\sqrt{2}}{\sqrt{3}}u_{-1},\frac{1}{\sqrt{3}}u_1-\frac{i\cdot\sqrt{2}}{\sqrt{3}}u_{-1}\right)\\ &= \hbar\frac{2\pi}{L}\cdot\left[\frac{1}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}}-\overline{\frac{i\cdot\sqrt{2}}{\sqrt{3}}}\cdot\frac{i\cdot\sqrt{2}}{\sqrt{3}}\right]\\ &= -\frac{1}{3}\cdot\hbar\frac{2\pi}{L}\end{aligned}$$

• We can, alternatively, compute the expectation value by using the outcomes of the measurement and their probabilities:

$$\begin{aligned}<p>|_\psi &= W_\psi(\hbar\frac{2\pi}{L})\cdot\hbar\frac{2\pi}{L}+W_\psi(-\hbar\frac{2\pi}{L})\cdot(-\hbar\frac{2\pi}{L})\\ &= \frac{1}{3}\cdot\hbar\frac{2\pi}{L}+\frac{2}{3}\cdot(-\hbar\frac{2\pi}{L})\\ &= -\frac{1}{3}\cdot\hbar\frac{2\pi}{L}\end{aligned}$$

• The state is a superposition of plane waves. These are also eigenstates of the Hamilton operator. Therefore we can read out the possible outcomes of a measurement of the energy by applying the Hamilton operator onto these states. For example:

$$\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\frac{1}{\sqrt{L}}e^{i\frac{2\pi}{L}\cdot x}=\frac{\hbar^2}{2m}\cdot(\frac{2\pi}{L})^2\frac{1}{\sqrt{L}}e^{i\frac{2\pi}{L}\cdot x}$$

Both plane waves carry the same energy eigenvalue. Accordingly, we have one possible outcome:

$$\frac{\hbar^2}{2m}\cdot(\frac{2\pi}{L})^2$$

• The probability of measuring this eigenvalue is "1".

•. The expectation value of the Hamilton operator is

$$\frac{\hbar^2}{2m}\cdot(\frac{2\pi}{L})^2$$

11. The parity operator.

The parity operator $\hat{P}$ is defined is one dimension as the operator that changes the sign of the spatial variable of a wave function:

$$\hat{P}\psi(x)=\psi(-x)$$

In particle physics it plays an important role – in particular because the weak interaction violates parity conservation (this means that when a process involves a weak interaction it proceeds in a different way in opposite spatial directions). It is therefore interesting to study its properties.

1. Show by explicit computation of the integral that $\hat{P}$ is an Hermitic operator, i.e.
$$\left(\hat{P}\varphi,\psi\right)=\left(\varphi,\hat{P}\psi\right)$$
2. Find all its possible eigenvalues (there are actually only two). Hint: solve the eigenvalue equation by using the fact that applying $\hat{P}$ twice gives the identity operation, i.e. $(\hat{P})^2=\mathbb{1}$.
3. In the space defined by the set $\{u_n(x)\}$, $n=0,\pm1,\pm2,....$ (see Exercise 10) $\hat{P}$ has eigenfunctions. Find all of them and establish which ones belong to which eigenvalue.
4. Consider again the wave function of Exercise 10
$$\psi(x)=\frac{1}{\sqrt{3}}\frac{1}{\sqrt{L}}e^{i\frac{2\pi}{L}\cdot x}+\frac{i\cdot\sqrt{2}}{\sqrt{3}}\frac{1}{\sqrt{L}}e^{-i\frac{2\pi}{L}\cdot x}$$

and find what are the possible outcomes of measuring the parity of a system in the state ψ and with which probability these values are measured.

5. Compute the quantum mechanical expectation value of the operator $\hat{P}$ in the state ψ by two methods: first use the values you have just computed for the possible outcome of an experiment and their probability. Second, by computing explicitly $\left(\psi,\hat{P}\psi\right)$.

Solution.

1.
$$\begin{aligned}\left(\hat{P}\,\psi,\psi\right) &= \int_{-\infty}^{\infty}\overline{\psi(-x)}\cdot\psi(x)\,dx \underbrace{=}_{y=-x,dx=-dy} -\int_{\infty}^{-\infty}\overline{\psi(y)}\psi(-y)\,dy\\ &= \int_{-\infty}^{\infty}\overline{\psi(y)}\hat{P}\,\psi(y)\,dy\\ &= \left(\psi,\hat{P}\,\psi\right)\end{aligned}$$

2. The sought for eigenvalues are the solution of the equation
$$\hat{P}\psi(x)=\lambda\cdot\psi(x)$$
As
$$(\hat{P})^2\psi(x)=\psi(x)=(\lambda)^2\psi(x)$$
we obtain
$$\lambda^2=1\longleftrightarrow\lambda=\pm1$$

3. From
$$\hat{P}\psi(x)=\psi(-x)=1\cdot\psi(x)$$
we learn that the eigenfunctions to $\lambda=1$ are even with respect to change of sign. The eigenfunctions to $\lambda=-1$ are the uneven ones. The set of plane waves
$$\{u_n(x)\}\quad n=0,\pm1,\pm2,....$$
are neither even nor uneven, so that they are not eigenfunctions of $\hat{P}$. However, we can construct a set of eigenfunctions for $\hat{P}$ by guessing suitable linear combinations:
$$v_n(x)=\frac{1}{\sqrt{2}}(u_n(x)+u_{-n}(x))=\sqrt{\frac{2}{L}}\cos\left(\frac{2\pi n}{L}x\right)\quad\hat{P}\,v_n=1\cdot v_n$$
$$w_n(x)=\frac{1}{\sqrt{2}}(u_n(x)-u_{-n}(x))=i\cdot\sqrt{\frac{2}{\cdot L}}\sin\left(\frac{2\pi n}{L}x\right)\quad\hat{P}\,w_n=(-1)\cdot w_n$$
We have taken care that the set of eigenfunctions of $\hat{P}$ are normalized to ”1”, for further use.

4. To read out from a given wave function the possible outcome of an experiment aimed at detecting the parity, it is necessary to expand the given wave function as a linear combinations of **eigenfunctions** of $\hat{P}$. We use
$$u_n=\frac{1}{\sqrt{2}}(v_n+w_n)\quad u_{-n}=\frac{1}{\sqrt{2}}(v_n-w_n)$$
to write
$$\psi=\frac{1}{\sqrt{6}}(1+i\sqrt{2})v_1+\frac{1}{\sqrt{6}}(1-i\sqrt{2})w_1$$
or
$$\psi(x)=\frac{1}{\sqrt{6}}(1+i\sqrt{2})\sqrt{\frac{2}{L}}cos\left(\frac{2\pi n}{L}x\right)+\frac{1}{\sqrt{6}}(1-i\sqrt{2})\sqrt{\frac{2}{L}}isin\left(\frac{2\pi n}{L}x\right)$$

We can now read out that a possible outcome of measuring the parity is the eigenvalue $+1$ with probability

$$\begin{aligned} W_\psi(+1) &= |<v_1,\psi>|^2 = |<v_1, \frac{1}{\sqrt{6}}(1+i\sqrt{2})v_1 + \frac{1}{\sqrt{6}}(1-i\sqrt{2})w_1>|^2 \\ &= |\frac{1}{\sqrt{6}}(1+i\sqrt{2})|^2 = \frac{1}{6}(1+i\sqrt{2})(1-i\sqrt{2}) \\ &= 1/2 \end{aligned}$$

and the eigenvalue -1 also with probability

$$W_\psi(+1) = 1/2$$

5.

$$<\hat{P}>|_\psi = \frac{1}{2}\cdot(+1) + \frac{1}{2}\cdot(-1) = 0$$

or (with $\frac{1}{\sqrt{6}}(1+i\sqrt{2}) \doteq A$ and $\frac{1}{\sqrt{6}}(1-i\sqrt{2}) \doteq B$)

$$\hat{P}\,\psi(x) = \hat{P}(Av_1 + Bw_1) = A\,v_1 - B\,w_1$$

and

$$\begin{aligned} \left(\psi, \hat{P}\,\psi\right) &= <Av_1 + Bw_1, A\,v_1 - Bw_1> = <Av_1, Av_1> + <Bw_1, -Bw_1> \\ &= |A|^2 - |B|^2 \\ &= 0 \end{aligned}$$

12. Self-adjoint operators.

DEFINITION. The operator that solves the equation

$$\left(L^\dagger g, f\right) = (g, Lf)$$

is called the adjoint operator to L.

DEFINITION. If $L^\dagger$ is formally identical to L AND the domain of definition of $(L^\dagger)$ coincide with the domain of definition of L, then the operator is said to be self-adjoint.

•Example 1: Study the momentum operator $-i\frac{d}{dx}$ in the interval $[a,b]$.

• Example 2: study the kinetic energy operator $H_0 \doteq -\frac{d^2}{dx^2}$ in the interval $[a,b]$**.**

Boundary conditions:

a. No boundary conditions at a and b.

b. Dirichlet boundary conditions.

c. Generalized periodic boundary conditions.

Solution.

• Example 1. We compute $p^\dagger$ using integration by parts:

$$\begin{aligned} (g, pf) &= -i\left[f(b)\overline{g(b)} - f(a)\overline{g(a)}\right] + \left(-i\frac{dg}{dx}, f\right) \\ &= \left(p^\dagger g, f\right) \end{aligned}$$

a. No boundary conditions at a and b. Because of the boundary term, $p^\dagger \neq p$.

b. Dirichlet boundary conditions. If we require $f(a) = f(b) = 0$ the boundary term evaluates to 0 and $p^\dagger = -i\frac{d}{dx} = p$, i.e. p is certainly Hermitic. However, the boundary terms vanishes even if $g(a) \neq g(b) \neq 0$, i.e. $D(g)$ is a larger set that $D(f)$. With these boundary conditions, p is not self-adjoint.

c. Generalized periodic boundary conditions. We require now

$$f(b) = f(a)\cdot e^{i\theta}$$

with some real θ (the case $\theta=0$ is called Born-von Karman boundary conditions). With these boundary conditions, the boundary terms become

$$f(a)\cdot e^{i\theta}\overline{g(b)}-f(a)\overline{g(a)}$$

In order for this boundary term to vanish we must have

$$g(b)=e^{i\theta}g(a)$$

i.e. the set of functions in the domain of $p^\dagger$ must obey the same boundary conditions as the set of functions in the domain of p. With these boundary conditions, p is self-adjoint.

• Example 2. Using partial integration we obtain

$$\begin{aligned}(g,H_0f)&=\overline{g(a)}f'(a)-\overline{g(b)}f'(b)+\overline{g'(b)}f(b)-\overline{g'(a)}f(a)+\left(-\frac{d^2g}{dx^2},f\right)\\&=\left(H_0^\dagger,f\right)\end{aligned}$$

a. No boundary conditions at a and b. Because of the boundary term, $H_0^\dagger\neq H_0$.

b. Dirichlet boundary conditions. Two boundary terms vanish because $f(a)=f(b)=0$. For the remaining two to vanish, $g(a)=g(b)=0$, i.e. the domain of definition for $H_0^\dagger$ and H_0 are identical and the operator H_0 is self-adjoint.

c. generalized periodic boundary conditions. We assume generalized periodic boundary conditions for f AND f':

$$\begin{aligned}f(b)&=f(a)e^{i\theta}\\f'(b)&=f'(a)e^{i\theta}\end{aligned}$$

The boundary terms write

$$\begin{aligned}&\overline{g(a)}f'(a)-\overline{g(b)}f'(a)\cdot e^{i\theta}+\overline{g'(b)}f(a)\cdot e^{i\theta}-\overline{g'(a)}f(a)\\&=f'(a)\left(\overline{g(a)}-\overline{g(b)}\cdot e^{i\theta}\right)-f(a)\left(\overline{g'(a)}-\overline{g'(b)}\cdot e^{i\theta}\right)\end{aligned}$$

The boundary term only vanishes if exactly the same boundary conditions are imposed on the g-function, so that these boundary conditions render H_0 self-adjoint.

13. The Hamilton operator $H=-\frac{d^2}{dx^2}+V(x)$ in the intervals $[0,\infty]$ and $[a,b]$.

Discuss the self-adjointness of this operator.

Solution.

We start with the interval $[0,\infty]$ and boundary condition $f'(0)=a\cdot f(0)$:

$$\begin{aligned}(g,Hf)&=\overline{g(0)}f'(0)-\overline{g'(0)}f(0)-\left(\overline{g(x)}f'(x)+\overline{g'(x)}f(x)\right)|_{x\to\infty}+\left(-\frac{d^2g}{dx^2}+V(x)g,f\right)\\&=\left(H^\dagger g,f\right)\end{aligned}$$

Provided all functions are square integrable, all boundary terms vanish if g obey the same boundary condition at $x=0$. With this boundary condition, H is self-adjoint.

In case of the interval $[a,b]$, we have:

$$\begin{aligned}\left(g,-\frac{d^2f}{dx^2}+V(x)f\right)&=\overline{g(a)}f'(a)-\overline{g'(a)}f(a)-\overline{g(b)}f'(b)+\overline{g'(b)}f(b)\\&+\left(-\frac{d^2g}{dx^2},f\right)+(V(x)g,f)\end{aligned}$$

– Dirichlet boundary condition $f(a)=g(b)=0$. The boundary term vanishes if $g(a)=g(b)=0$, i.e. the operator is self-adjoint. – von Neumann boundary condition $f'(a)=f'(b)=0$. The boundary

term vanishes if also $g'(a) = g'(b) = 0$, i.e. the Hamilton operator is self-adjoint.
– Generalized periodic boundary conditions. In the case of generalized periodic boundary conditions

$$f(a) = g(f)e^{i\theta} \quad f'(a) = f'(b)e^{i\theta}$$

the boundary term writes

$$f'(b)\left(\overline{g(a)}e^{i\theta} - \overline{g(b)}\right) + f(b)\left(\overline{g'(b)}e^{i\theta} - \overline{g'(a)}\right)$$

and vanishes if g also fulfills the conditions $g(a) = g(b)e^{i\theta}$ and $g'(a) = g'(b)e^{i\theta}$, i.e. the operator is self-adjoint.

12.3. Exercises to Chapter 5.

1. Spin $\frac{1}{2}$ algebra.

The matrix representation of the x-, y- und z-components of the Pauli operators reads:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad , \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad .$$

- Compute explicitely the matrix representation of σ^2
- Show that the vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} ; \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

are eigenvectors of σ_z and σ^2.

- Consider the rules of computation for Pauli matrices:

$$\sigma_x Y^{\pm} = Y^{\mp} \quad \sigma_y Y^{\pm} = \pm i Y^{\mp} \quad \sigma_z Y^{\pm} = \pm Y^{\pm} \quad \sigma \cdot \sigma Y^{\pm} = 3Y^{\pm}$$

(you have just demonstrated two of these rules). Show that

$$\sqrt{\frac{1}{2}}(Y^{+} \otimes Y^{-} - Y^{-} \otimes Y^{+})$$

(the singlet spin state) is indeed an eigenstate of the operator

$$(\sigma \otimes \mathbb{1} + \mathbb{1} \otimes \sigma)^2$$

of the composite system of two spin $\frac{1}{2}$-particles to the total spin quantum number $S = 0$.

Solution.

-

$$\begin{aligned}
\sigma^2 &= \sigma_x \cdot \sigma_x + \sigma_y \cdot \sigma_y + \sigma_z \cdot \sigma_z \\
&= \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right\} \\
&= \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \\
&= 3 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}$$

-

$$\begin{pmatrix} 1 & 0 \\ 0 & \bar{1} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & \bar{1} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \bar{1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$3\cdot\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{pmatrix}1\\0\end{pmatrix}=3\begin{pmatrix}1\\0\end{pmatrix}\quad 3\cdot\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{pmatrix}0\\1\end{pmatrix}=3\cdot\begin{pmatrix}0\\1\end{pmatrix}$$

-

$$\begin{aligned}(\sigma\otimes\mathbb{1}+\mathbb{1}\otimes\sigma)^2 &= (\sigma\otimes\mathbb{1})^2+(\mathbb{1}\otimes\sigma)^2+2\sigma\otimes\sigma\\ &= \sigma\cdot\sigma\otimes\mathbb{1}+\mathbb{1}\otimes\sigma\cdot\sigma+2\sigma\otimes\sigma\\ &= 6\mathbb{1}\otimes\mathbb{1}+2\sigma\otimes\sigma\end{aligned}$$

$$\begin{aligned}[6\mathbb{1}\otimes\mathbb{1}+2\sigma\otimes\sigma]\sqrt{\frac{1}{2}}(Y^+\otimes Y^- - Y^-\otimes Y^+) &= 6\cdot\sqrt{\frac{1}{2}}(Y^+\otimes Y^- - Y^-\otimes Y^+)\\ &+ 2[\sigma_x\otimes\sigma_x]\sqrt{\frac{1}{2}}(Y^+\otimes Y^- - Y^-\otimes Y^+)\\ &+ 2[\sigma_y\otimes\sigma_y]\sqrt{\frac{1}{2}}(Y^+\otimes Y^- - Y^-\otimes Y^+)\\ &+ 2[\sigma_z\otimes\sigma_z]\sqrt{\frac{1}{2}}(Y^+\otimes Y^- - Y^-\otimes Y^+)\\ &= 6\cdot\sqrt{\frac{1}{2}}(Y^+\otimes Y^- - Y^-\otimes Y^+)\\ &+ 2(\sqrt{\frac{1}{2}}(Y^-\otimes Y^+ - Y^+\otimes Y^-)\\ &+ 2\cdot(i\cdot\bar{i})(\sqrt{\frac{1}{2}}(Y^-\otimes Y^+ - Y^+\otimes Y^-)\\ &+ 2\cdot(1\cdot\bar{1})(\sqrt{\frac{1}{2}}(Y^+\otimes Y^- - Y^-\otimes Y^+)\\ &= 0\cdot(\sqrt{\frac{1}{2}}(Y^+\otimes Y^- - Y^-\otimes Y^+)\end{aligned}$$

2. Algebra of angular momentum for composite systems.

Given is a composite system consisting of two particles with angular momentum quantum numbers J_1 and J_2.

• Show by using the rules of computation for angular momentum operators that the $(2J_1+1)\cdot(2J_2+1)$ product states

$$\{Z_{J_1}^{m_1}\otimes Z_{J_2}^{m_2}\}$$

are eigenfunctions of the operators $J_{1,z}$, $J_{2,z}$, $\mathbf{J}_1^2$ und $\mathbf{J}_2^2$. Find the corresponding eigenvalues and degeneracy.

Remark: the states

$$\{Z_J^m\}$$

obey the relations

$$J_z Z_J^m=\hbar\cdot m\cdot Z_J^m \quad (\mathbf{J})^2 Z_J^m=\hbar^2\cdot J(J+1)Z_J^m \quad J=0,\frac{1}{2},1,.....$$

Solution.

1.

$$(J_{1,z}\otimes\mathbb{1})Z_{J_1}^{m_1}\otimes Z_{J_2}^{m_2}=\hbar\cdot m_1\cdot Z_{J_1}^{m_1}\otimes Z_{J_2}^{m_2}$$

Eigenvalue: $\hbar \cdot m_1$, $2J_2+1$-times degenerate.

2. $$(\mathbb{1} \otimes J_{2,z}) Z_{J_1}^{m_1} \otimes Z_{J_2}^{m_2} = \hbar \cdot m_2 \cdot Z_{J_1}^{m_1} \otimes Z_{J_2}^{m_2}$$
   Eigenvalue: $\hbar \cdot m_2$, $2J_1+1$-times degenerate.

3. $$(\mathbf{J}_1^2 \otimes \mathbb{1}) Z_{J_1}^{m_1} \otimes Z_{J_2}^{m_2} = \hbar^2 J_1(J_1+1) \cdot Z_{J_1}^{m_1} \otimes Z_{J_2}^{m_2}$$
   Eigenvalue: $\hbar^2 \cdot J_1(J_1+1)$, $(2J_1+1)\cdot(2J_2+1)$-times degenerate.

4. $$(\mathbb{1} \otimes \mathbf{J}_2^2) Z_{J_1}^{m_1} \otimes Z_{J_2}^{m_2} = \hbar^2 J_2(J_2+1) \cdot Z_{J_1}^{m_1} \otimes Z_{J_2}^{m_2}$$
   Eigenvalue: $\hbar^2 \cdot J_2(J_2+1)$, $(2J_1+1)\cdot(2J_2+1)$-times degenerate.

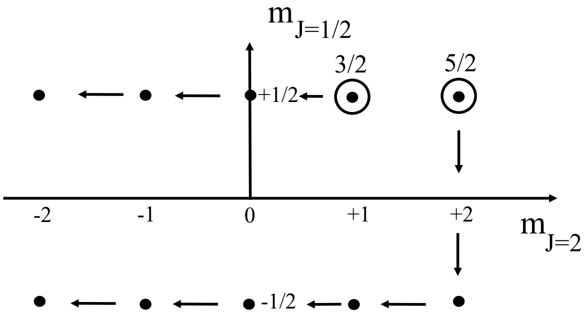
3. Cartan diagram.

Consider a composite system consisting of one particle with angular momentum quantum number 2 and one particle with angular momentum quantum number $\frac{1}{2}$.

1. Build the corresponding Cartan lattice and find the possible eigenvalues of $\hat{J}_{1z}+\hat{J}_{2z}$ as Cartan weights (set $\hbar = 1$ for simplicity).
2. Encircle on the Cartan diagram the possible quantum numbers for the operator $(\hat{\mathbf{J}}_1+\hat{\mathbf{J}}_2)^2$.
3. Find the possible eigenvalues for the operator $(\hat{\mathbf{J}}_1+\hat{\mathbf{J}}_2)^2$.
4. Write all possible eigenstates of $\hat{J}_{1z}+\hat{J}_{2z}$
5. Write the matrix representation of $\hat{J}_{1z}+\hat{J}_{2z}$ in this set of basis functions.
6. The matrix representation of $(\hat{\mathbf{J}}_1+\hat{\mathbf{J}}_2)^2$ in this basis set is not diagonal. However, there are special linear combinations that diagonalize the matrix. How does the matrix representation of $(\hat{\mathbf{J}}_1+\hat{\mathbf{J}}_2)^2$ look like after diagonalization?

Solution.

1. The Cartan weigths read: $\frac{5}{2}, \frac{3}{2}, \frac{3}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}, -\frac{3}{2}, -\frac{5}{2}$



Cartan diagram for $\mathbf{2} \otimes \mathbb{1} + \mathbb{1} \otimes \frac{1}{2}$. The dots are the Cartan lattice. The Cartan weigths indicating the possible values for the total angular momentum quantum numbers are encircled and given in the diagram: follow the arrows to sample their z-components.

2. The possible quantum numbers for $(\hat{\mathbf{J}}_1+\hat{\mathbf{J}}_2)^2$ are
$$J=\frac{5}{2}\quad J=\frac{3}{3}$$

3. The eigenvalues write
$$\frac{35}{4}\quad\frac{15}{4}$$

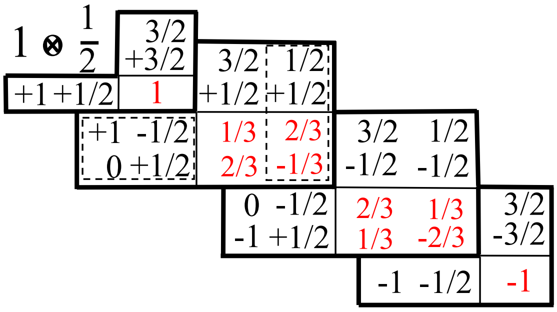
4. There are 10 product states
$$Z_2^{m_2}\otimes Z_{\frac{1}{2}}^{m_{\frac{1}{2}}}$$
The values of m_2 and $m_{\frac{1}{2}}$ are along the horizontal respectively vertical axis of the Cartan diagram.

5. Let us use now $\hbar$:
$$J_z=\begin{pmatrix}
\frac{5}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & \frac{3}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & \frac{3}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{2} & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{2} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{5}{2}
\end{pmatrix}$$

6. Solution:
$$(\mathbf{J})^2=\hbar^2\begin{pmatrix}
\frac{35}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & \frac{35}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & \frac{35}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & \frac{35}{4} & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & \frac{35}{4} & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & \frac{35}{4} & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & \frac{15}{4} & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{15}{4} & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{15}{4} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{15}{4}
\end{pmatrix}$$

4. Clebsch-Gordan coefficients.

The Table for the Clebsch-Gordan coefficients (taken from M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018)), p.564) related to an electron with orbital angular momentum quantum number ”1” and spin quantum number ”$\frac{1}{2}$” is given in the next figure.



$1\otimes\frac{1}{2}$	3/2 +3/2
+1 +1/2	1

	3/2 +1/2	1/2 +1/2
+1 -1/2	1/3	2/3
0 +1/2	2/3	-1/3

	3/2 -1/2	1/2 -1/2
0 -1/2	2/3	1/3
-1 +1/2	1/3	-2/3

	3/2 -3/2
-1 -1/2	-1

Table of the Clebsch-Gordan coefficients (in red) for $\mathbf{1}\oplus\frac{1}{2}$.

1. Obtain from the Table the two eigenstates of the total angular momentum $(\mathbf{L}\otimes\mathbb{1}+\mathbb{1}\otimes\mathbf{S})^2$ to the total angular momentum quantum number $J=\frac{1}{2}$ call them $|\, a>$ and $|\, b>$.
2. What is the outcome of an experiment that measures the square of the total angular momentum $\mathbf{J}^2=(\mathbf{L}+\mathbf{S})^2$ in $|\, a>$ and $|\, b>$?
3. What is the outcome of an experiment that measures the z-component of the total angular momentum $J_z \doteq L_z+S_z$ in the states $|\, a>$ and $|\, b>$?
4. Compute the matrix representation of J_z in the subspace built by the two states $|\, a>$ and $|\, b>$.
5. What is the outcome of an experiment that measures the z-component of the orbital angular momentum L_z in the states $|\, a>$ and $|\, b>$? What are the amplitudes for the respective outcomes?
6. Compute the expectation value of L_z is the states $|\, a>$ and $|\, b>$.
7. Compute the matrix representation of L_z in the subspace built by the two states $|\, a>$ and $|\, b>$ (the matrix is diagonal)
8. Show that, despite the matrix being diagonal, $|\, a>$ and $|\, b>$ are NOT eigenfunctions of L_z.
9. Compute the matrix representation of S_z in the subspace built by the two states $|\, a>$ and $|\, b>$ (the matrix is diagonal).
10. Show that the matrix of L_z added to the matrix of S_z gives indeed the matrix of J_z.
11. Show that, despite the matrix being diagonal, $|\, a>$ and $|\, b>$ are NOT eigenfunctions of S_z.

Solution.

1. The eigenstate to $J=\frac{1}{2}$ and $J_z=+\frac{1}{2}$ is found by using the Clebsch-Gordan coefficients that appear in the intersection of the two dashed rectangles:
$$|\,\frac{1}{2},\frac{1}{2}>\doteq|\, a>\sqrt{\frac{2}{3}}Y_1^1\otimes Y^- - \sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+$$
The second eigenstate is
$$|\,\frac{1}{2},-\frac{1}{2}>|\, b>=\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^- - \sqrt{\frac{2}{3}}Y_1^{-1}\otimes Y^+$$

2. In both states, the outcome is
$$J(J+1)\hbar^2 = \hbar^2 \cdot \frac{3}{4}$$
3. From the Table, we read out that $| a >$ is an eigenstate of J_z to the eigenvalue
$$\hbar \frac{1}{2}$$
and $| b >$ is an eigenstate to the eigenvalue
$$-\hbar \frac{1}{2}$$
so that the outcome of a measurement of J_z is $\hbar\frac{1}{2}$ in $| a >$ and $-\frac{1}{2}\hbar$ in $| b >$.
4. We know that $| a >$ and $| b >$ are eigenstates of the operator J_z which means that the matrix representation of J_z in this basis is diagonal with the eigenvalues of $| a >$ and $| b >$ on the diagonal:
$$\hbar \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$
5. In $| a >$ one will detect the value
$$0 \cdot \hbar \quad \text{amplitude} \quad -\sqrt{\frac{1}{3}}$$
$$\hbar \quad \text{amplitude} \quad +\sqrt{\frac{2}{3}}$$
In $| b >$ one will detect the value
$$0 \quad \text{amplitude} \quad \sqrt{\frac{1}{3}}$$
$$-\hbar \quad \text{amplitude} \quad -\sqrt{\frac{2}{3}}$$
6.
$$< L_z >_a = 0 \cdot \hbar \cdot (-\sqrt{\frac{1}{3}})^2 + \hbar \cdot (\sqrt{\frac{2}{3}})^2 = \frac{2}{3} \cdot \hbar$$
$$< L_z >_b = 0 \cdot \hbar \cdot (\sqrt{\frac{1}{3}})^2 - \hbar \cdot (\sqrt{\frac{2}{3}})^2 = -\frac{2}{3} \cdot \hbar$$
7.
$$\begin{aligned} < a | L_z | a > &= \left(-\sqrt{\frac{1}{3}} \cdot Y_1^0 \otimes Y^+ + \sqrt{\frac{2}{3}} Y_1^1 \otimes Y^-, \underbrace{L_z(-\sqrt{\frac{1}{3}} \cdot Y_1^0 \otimes Y^+ + \sqrt{\frac{2}{3}} Y_1^1 \otimes Y^-)}_{\sqrt{\frac{2}{3}} \cdot \hbar Y_1^1 \otimes Y^-} \right) \\ &= \left(\sqrt{\frac{2}{3}} Y_1^1 \otimes Y^-, \sqrt{\frac{2}{3}} \cdot \hbar Y_1^1 \otimes Y^- \right) \\ &= = \hbar \cdot \frac{2}{3} \end{aligned}$$
Similarly, one finds that
$$< b | L_z | b > = -\hbar \cdot \frac{2}{3}$$

The non-diagonal matrix elements are zero so that the matrix representation of L_z in this subspaces writes

$$\hbar\begin{pmatrix} \frac{2}{3} & 0 \\ 0 & -\frac{2}{3} \end{pmatrix}$$

8.
$$\begin{aligned} L_z(-\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+ + \sqrt{\frac{2}{3}}Y_1^1\otimes Y^-) &= L_z(-\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+) + L_z(\sqrt{\frac{2}{3}}Y_1^1\otimes Y^-) \\ &= -\sqrt{\frac{1}{3}}\cdot\underbrace{(L_z Y^0)}_{0}\otimes Y^+ + \sqrt{\frac{2}{3}}\cdot\underbrace{(L_z Y^1)}_{\hbar\cdot 1}\otimes Y^- \\ &= \sqrt{\frac{2}{3}}\cdot\hbar Y_1^1\otimes Y^- \not\propto (-\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+ + \sqrt{\frac{2}{3}}Y_1^1\otimes Y^-) \end{aligned}$$

9. The top left matrix element reads
$$\begin{aligned} <a\,|\,S_z\,|\,a> &= \left(-\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+ + \sqrt{\frac{2}{3}}Y_1^1\otimes Y^-, \underbrace{S_z(-\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+ + \sqrt{\frac{2}{3}}Y_1^1\otimes Y^-)}_{\frac{\hbar}{2}\cdot(-\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+ - \sqrt{\frac{2}{3}}Y_1^1\otimes Y^-)}\right) \\ &= -\frac{\hbar}{6} \end{aligned}$$

The second diagonal matrix elements is $\frac{\hbar}{6}$, the non diagonals are vanishing.

$$S_z = \begin{pmatrix} -\frac{\hbar}{6} & 0 \\ 0 & \frac{\hbar}{6} \end{pmatrix}$$

10. The matrix of L_z plus the matrix of S_z should give the matrix of J_z in this subspace:
$$\begin{pmatrix} \frac{2\hbar}{3} & 0 \\ 0 & -\frac{2\hbar}{3} \end{pmatrix} + \begin{pmatrix} -\frac{\hbar}{6} & 0 \\ 0 & \frac{\hbar}{6} \end{pmatrix} = \begin{pmatrix} \frac{\hbar}{2} & 0 \\ 0 & -\frac{\hbar}{2} \end{pmatrix}$$

11.
$$\begin{aligned} S_z(-\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+ + \sqrt{\frac{2}{3}}Y_1^1\otimes Y^-) &= \frac{\hbar}{2}\cdot(-\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+ - \sqrt{\frac{2}{3}}Y_1^1\otimes Y^-) \\ \not\propto (-\sqrt{\frac{1}{3}}\cdot Y_1^0\otimes Y^+ + \sqrt{\frac{2}{3}}Y_1^1\otimes Y^-) \end{aligned}$$

5. Eigenvectors of the $\sigma_1\cdot\sigma_2$ operator.

This operator for the composite system of two-spin $\frac{1}{2}$ particles is defined, precisely speaking, as follows:

$$\sigma_{1x}\otimes\sigma_{2x} + \sigma_{1y}\otimes\sigma_{2y} + \sigma_{1z}\otimes\sigma_{2z}$$

In the space of the four product states

$$Y^+Y^+ \quad Y^+Y^- \quad Y^-Y^+ \quad Y^-Y^-$$

find the eigenvalues and the eigenvectors of the operator.

Solution.

We build the four-by-four matrix representing this operator in this subspace and diagonalize it. We need only 10 matrix elements, as the operator is hermitic. We compute using the rules of computation

$$\sigma_x Y^{\pm} = Y^{\mp} \quad \sigma_y Y^{\pm} = \pm i Y^{\mp} \quad \sigma_z Y^{\pm} = \pm Y^{\pm} \quad \sigma \cdot \sigma Y^{\pm} = 3Y^{\pm}$$

For example:

$$\begin{aligned} (\sigma_1 \otimes \sigma_2)_{23} &= (Y^+Y^-, \sigma_x \otimes \sigma_x Y^- Y^+) \\ &+ (Y^+Y^-, \sigma_y \otimes \sigma_y Y^- Y^+) \\ &+ (Y^+Y^-, \sigma_z \otimes \sigma_z Y^- Y^+) \\ &= 1 \cdot 1 + (-i \cdot i) + 0 \cdot 0 \\ &= 2 \end{aligned}$$

We find the matrix representation

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bar{1} & 2 & 0 \\ 0 & 2 & \bar{1} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

One eigenvalue can be read out from the matrix directly, as the matrix is partially diagonal:

$$+1 \quad \text{twice degenerate}$$

The remaining eigenvalues are found by solving the determinantal equation of the central block of the matrix:

$$(-1-\lambda)^2 - 4 = 0 \quad \longleftrightarrow \lambda = 1 \quad \lambda = -3$$

Accordingly +1 is three times degenerate, −3 once degenerate. The eigenvectors are found by solving the set of coupled linear algebraic equations for the sought-for coeffcients x_1, x_2, x_3, x_4 (inserting the known eigenvalues):

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

The eigenvector to the eigenvalue −3 is (singlet)

$$0, 1, \bar{1}, 0 \quad \longleftrightarrow \sqrt{\frac{1}{2}} [Y^+(1)Y^-(2) - Y^-(1)Y^+(2)]$$

The three eigenvectots to +1 are (triplet)

$$1. \quad Y^+(1)Y^+(2)$$

$$2. \quad \sqrt{\frac{1}{2}} [Y^+(1)Y^-(2) + Y^-(1)Y^+(2)]$$

$$3. \quad Y^-(1)Y^-(2)$$

12.4. Exercises to Chapter 7, 8.

1. Multipletts, terms and the Zeeman effect in the Si-atom.

The Si atom has 14 electrons.

1. What is the electron configuration of the ground state?
2. What is the spectral symbol of the ground state?

3. What is the outcome of measuring the square of the vector of the total angular momentum in the ground state?

4. What is the outcome of measuring the square of the vector of the orbital angular momentum in the ground state?

5. What is the outcome of measuring the square of the vector of the spin angular momentum in the ground state?

6. How much is the degeneracy of the ground state?

7. One electron from the outer shell is removed and only one is left in the outer shell. Which multiplet ${}^{2S+1}L$ can be built in this one electron configuration?

8. Write the two spectral **terms** that are obtained from the remaining multiplet.

9. Which spectral term is the excited state?

10. How much is the degeneracy of the excited state?

11. Find the four basis eigenfunctions in the excited state from the table of Clebsch-Gordan coefficients.

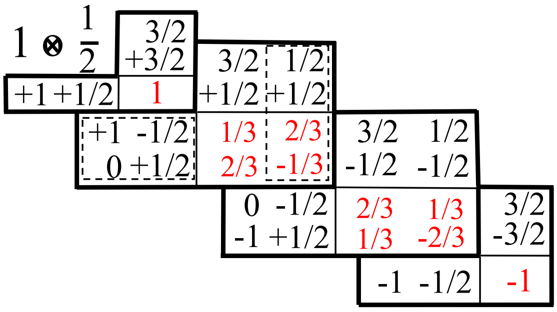


Table of the Clebsch-Gordan coefficients (in red) for $\mathbf{1} \oplus \frac{1}{2}$.

12. Construct the matrix representation of $L_z + 2 \cdot S_z$ in this space (note that the matrices for L_z and S_z are diagonal.)

13. Obtain from this matrix the matrix representation of the Zeeman operator, assuming a magnetic field of strength B is applied along $+z$.

14. Compute the Landé factor for this spectral term

15. Show that the matrix representation of the Zeeman operator obtained using the Landé factor is identical to the matrix representation you have obtained by summing L_z and $2S_z$.

16. Find all possible eigenvalues of the Zeeman operator.

17. Which eigenvalue describes the state with highest energy?

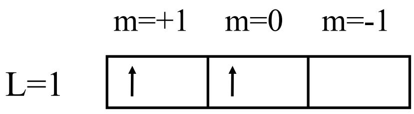
18. Which one is its eigenfunction?

19. What is the outcome of an experiment that measures the value of $\mathbf{J}^2$ in this state?

20. What is the outcome of an experiment that measures the value of J_z in this state?
21. What is the outcome of an experiment that measures the value of μ_z (the z-component of the magnetic moment) in this state?

Solution.

1. The electronic configuration writes
$$(1s)^2(2s)^2(2p)^6(3s)^2(3p)^2$$
2. To find the spectral symbol of the ground state we need to apply the three Hund's rules. The next figure shows the filling of the states according to the Hund's rules.



The two electrons are placed both with spin up in the orbitals with $m = +1$ and $m = 0$, thus maximizing both the spin and the orbital angular momentum.

   The total spin is $S = 1$ and the total orbital angular momentum is $L = 1+0 = 1$, i.e. the ground state multiplett of the Si-atom is ^3P. There are three possible values for the total angular momentum quantum number: $J = 2, 1, 0$ The third's Hund rule prescribes that, by less than half filling, the total orbital momentum quantum number is $\mid L - S \mid$, in this case $J = 0$. The spectroscopic symbol of the ground state is
$$^3P_0$$
3. Outcome for $\mathbf{J}^2$: $J(J+1)\cdot\hbar^2 = 0$.
4. Outcome for $\mathbf{L}^2$: $L(L+1)\hbar^2 = 2\hbar^2$.
5. Outcome for $\mathbf{S}^2$: $S(S+1)\hbar^2 = 2\hbar^2$.
6. The degeneracy of the ground state belonging to $J = 0$ is given by $2J + 1 = 1$.
7. If one electron is removed, the multiplett is changed to ^2P, with $S = \frac{1}{2}$ and $L = 1$.
8. One can obtain two spectral terms by using $J = |L - S|$ and $J = |L + S|$. The first one describes the ground state and the second one describes an excited state: $^2P_{1/2}$ and $^2P_{3/2}$.
9. The excited state is described by the spectral term which violates the third Hund's rule: $^2P_{3/2}$ (the spectral term $^2P_{1/2}$ describes the ground state).
10. The degeneracy is given by $2J + 1 = 4$.

11. We find the wavefunctions for $^2P_{3/2}$ if we use the columns with heading "$J=3/2$":

$$\begin{aligned}
\psi_1 &= Y_1^1 Y^+ \\
\psi_2 &= \sqrt{\frac{1}{3}} Y_1^1 Y^- + \sqrt{\frac{2}{3}} Y_1^0 Y^+ \\
\psi_3 &= \sqrt{\frac{2}{3}} Y_1^0 Y^- + \sqrt{\frac{1}{3}} Y_1^{-1} Y^+ \\
\psi_4 &= Y_1^{-1} Y^-
\end{aligned}$$

12.

$$\begin{aligned}
<\psi_1,(L_z+2S_z)\psi_1> &= <Y_1^1 Y^+,(L_z+2S_z)Y_1^1 Y^+> \\
&= 2\hbar <Y_1^1 Y^+, Y_1^1 Y^+> \\
&= 2\hbar \\
<\psi_2,(L_z+2S_z)\psi_2> &= \frac{2}{3}\hbar \\
<\psi_3,(L_z+2S_z)\psi_3> &= -\frac{2}{3}\hbar \\
<\psi_4,(L_z+2S_z)\psi_4> &= -2\hbar
\end{aligned}$$

$$L_z+2S_z = \begin{pmatrix} 2\hbar & 0 & 0 & 0 \\ 0 & 2/3\hbar & 0 & 0 \\ 0 & 0 & -2/3\hbar & 0 \\ 0 & 0 & 0 & -2\hbar \end{pmatrix}$$

13.

$$\mathbf{H_Z} = \frac{\mu_B}{\hbar}(L_z+2S_z)B = \mu_B B \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 0 \\ 0 & 0 & -2/3 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

14.

$$g_{LSJ} = 1 + \frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)} = \frac{4}{3}$$

by using $J=\frac{3}{2}$, $S=\frac{1}{2}$ and $L=1$.

15.

$$\mathbf{H_Z} = \frac{4}{3}\mu_B B \begin{pmatrix} \frac{3}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{3}{2} \end{pmatrix} = \mu_B B \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 0 \\ 0 & 0 & -2/3 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} = \mu_B(L_z+2S_z)B.$$

16. The eigenvalues of the Zeeman operator van be read out from the diagonal of the matrix to be

$$\pm 2\mu_B B \quad \pm\frac{2}{3}\mu_B B$$

17. The highest energy is

$$E_{term} + 2\mu_B B$$

18. The eigenstate with highest energy is

$$\psi_1 = Y_1^1 Y^+$$

19. For the state ψ_1 the quantum number J is still $\frac{3}{2}$, which means that the outcome of measuring $\mathbf{J}^2$ is
$$J(J+1)\hbar = \frac{15}{4}\hbar$$

20. ψ_1 is the eigenstate with $J_z = \frac{3}{2}$, so that the outcome of measuring J_z is
$$\frac{3}{2}\hbar$$

21. The vector of the magnetic moment operator is defined as
$$\mu = -\mu_B \cdot g_{LSJ} \cdot \frac{\mathbf{J}}{\hbar}$$
Applying this definition to the highest J_z eigenvalue we obtain
$$\mu_z = -\frac{4}{3} \cdot \mu_B \cdot \frac{3}{2} = -2\mu_B$$

2. Multipletts, terms and the Zeeman effect in the Zr-atom.

We consider a Zr-Atom (40 electrons). This is one of the atoms where the 5s level if filled before the 4d levels.

1. Write the electron configuration for the ground state.
   Solution: $(Kr)(5s)^2(4d)^2$

2. Write the multiplet symbol of the ground state.
   Solution: ${}^3F$

3. Write the spectral symbol of the ground state.
   Solution: ${}^3F_2$

4. What is the total angular momentum quantum number of the ground state?
   Solution: 2

5. What is the eigenvalue of the square of the total angular momentum vector in the ground state?
   Solution: $\hbar^2 \cdot 6$

6. Write the matrix representation of the operator $(\mathbf{L}+\mathbf{S})^2$ in the eigenspace of the ground state.
   Solution:
$$\hbar^2 \cdot \begin{pmatrix} 6 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{pmatrix}$$

7. Write the matrix representation of the operator $(\mathbf{L})^2$ in the eigenspace of the ground state.
   Solution:
$$\hbar^2 \cdot \begin{pmatrix} 12 & 0 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 & 12 \end{pmatrix}$$

8. Write the matrix representation of the operator $(\mathbf{S})^2$ in the eigenspace of the ground state.
Solution:
$$\hbar^2\cdot\begin{pmatrix}2&0&0&0&0\\0&2&0&0&0\\0&0&2&0&0\\0&0&0&2&0\\0&0&0&0&2\end{pmatrix}$$

9. What are the possible outcomes for measuring the z-component of the total angular momentum vector in the ground state?
Solution: $\hbar\cdot\{2,1,0,\bar{1},\bar{2}\}$

10. A magnetic field with strength B along the z-direction is applied to the system. The energy level E_0 of the ground state splits into so called Zeeman sublevels. Compute the distance between two neighboring Zeeman levels as a function of B.

11. In the lowest lying Zeeman level, which value will be detected when the z-component of the total angular momentum vector is measured?

12. In the lowest lying Zeeman level, which value will be detected when the z-component of the magnetic moment is measured?

Solution of 10, 11 and 12.

10. The Zeeman splitting is given by
$$\mu_B\cdot g_{LSJ}\cdot B$$
We compute
$$g_{LSJ}=1+\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}=\frac{2}{3}$$
In this case, the Zeeman splitting amount to
$$\frac{2}{3}\cdot\mu_B\cdot B$$
11. The state with the lowest Zeeman energy is the one with
$$J_z=-2\cdot\hbar$$
12.
$$\mu_z=-\mu_B\cdot g_{LSJ}\cdot J_z=\frac{4}{3}\mu_B$$

3. One and two identical particles in quadratic potential.

We consider a spinless particle with mass m moving according to the Schrödinger equation in the xz-space. The potential energy is taken to be $\frac{K}{2}\cdot(x^2+z^2)$.

1. Write the Schrödinger equation that must be solved to determine the eigenstates to a given eigenvalue E, $\psi_E(x,z)$.
2. Find all possible energy eigenvalues, E, using your knowledge from the lecture, without solving equations. Notice: your expression for the energy must include the parameter K.
3. Which energy eigenvalue corresponds to the ground state state?
4. Find the energy value of the lowest lying EXCITED STATE.

5. What is the degeneracy of the lowest lying EXCITED STATE
6. Find the energy value of the next lowest lying EXCITED STATE.
7. What is the degeneracy of the next lowest lying EXCITED STATE?
8. The parity operator was defined in the lecture. Define it here again for this specific case.
9. What is the possible outcome of measuring the parity in the ground state? Your answer must be supported by an approximate hint about how the ground state wave function looks like.
10. Let us add a second, identical spinless particle with the same mass. The second particle also moves in the same potential energy and is not interacting with the first one. We have therefore two non-interacting particles with the same potential energy given by the expression above. Write the Schrödinger equation for the system consisting of the two particles.
11. Use your knowledge from the lecture to write the general expression for the energy eigenvalues of the two-particle system.
12. Let us give to the ground state the two-particle spinless system the lable $|\,0,0,0,0>$ (or $\psi_{0,0,0,0}$ is you like). Can you relate the four "0"'s to the energy of the ground state of the two-particle spinless system?
13. Write the (approximate) ground state wave function, i.e.
$$<....\,|\,0,0,0,0>$$
or, if you prefer
$$\psi_{0,0,0,0}(....)$$
14. Let us give now to each particle a spin $\frac{1}{2}$, but let us first ignore the Pauli principle. Write all basis states building the ground state, including the spin components.
15. "Switch on" the Pauli principle. Write again all basis states remaining in the ground state of the system.

Solution.

1.
$$-\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}+\right]\psi(x,z)+\frac{K}{2}\cdot(x^2+z^2)\psi(x,z)=E\psi(x,z)$$
2. Along both degrees of freedom x and z one recognizes the quantum mechanical harmonic oscillator. According to the rules of merging two non-interacting degrees of freedom, the energy eigenvalues of both degrees of freedom are added to produce the total eigenvalues. Accordingly
$$E(n_x,n_z)=\hbar\sqrt{\frac{K}{m}}\cdot(n_x+n_z+1)\quad n_i=0,1,2,3....$$
n_x, n_z are the quantum numbers necessary to lable the energy eigenvalues.
3. Set $n_x=n_z=0$
$$E=\hbar\sqrt{K}m\cdot 1$$
4. Set $n_x=0, n_z=1$ or $n_x=1, n_z=0$
$$E=\hbar\sqrt{K}m\cdot 2$$

5. The corresponding excited state is twice degenerate, as two possible choices of the quantum numbers give the same eigenvalue.

6. Set $n_x = n_z = 1$
$$E = \hbar\sqrt{K}m \cdot 3$$

7. This state is, as the ground state, once degenerate.

8. $$P\psi(x,z) \doteq \psi(-x,-z)$$

9. We remember: the ground state function of the one-dimensional harmonic oscillator is a gaussian centered at the origin of the quadratic potential. The wave function of the composite system is therefore
$$\propto e^{-[(\frac{x}{a})^2+(\frac{z}{a})^2]}$$
(a being some characteristic length, not relevant for discussing the parity). This function is even with respect to change of sign, so that the we will detect the eigenvalue +1 of the parity operator.

10. The SE for the two-particle system writes
$$-\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial z_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial z_2^2}\right]\psi(x_1,z_1,x_2,z_2)$$
$$+[\frac{K}{2}\cdot(x_1^2+z_1^2)+\frac{K}{2}\cdot(x_2^2+z_2^2)]\psi(x_1,z_1,x_2,z_2) = E\psi(x_1,z_1,x_2,z_2)$$

11. The the energy eigenvalues of the two non-interacting, identical particles are obtained by summing the energy eigenvalues:
$$E = \hbar\sqrt{\frac{K}{m}}\cdot(n_{x,1}+n_{x,2}+n_{z,1}+n_{z,2}+2) \quad n_i = 0,1,2,3....$$
Four quantum numbers n_i are necessary to label the energy eigenvalues: two pertaining the xz-degrees of freedom, two partaining the two-particles degrees of freedom.

12. The energy of the ground state of the two-particles system is obtained by inserting
$$n_{x,1} = n_{x,2} = n_{z,1} = n_{z,2} = 0$$
in the general expression for the eigenvalues. The four "0"s in the labeling of the ground state refer to the values of the quantum numbers.

13. The ground state wave function is the representation of the ground state with the variables (x_1,z_1,x_2,z_2):
$$< x_1,z_1,x_2,z_2 \mid 0,0,0,0 >= \psi_{0,0,0,0}(x_1,z_1,x_2,z_2) \propto e^{-[(\frac{x_1}{a})^2+(\frac{z_1}{a})^2]}\cdot e^{-[(\frac{x_2}{a})^2+(\frac{z_2}{a})^2]}$$

14. There are four eigenstates in the ground level:
$$\psi_{0,0,0,0}\otimes\begin{cases} Y^+Y^+ \\ Y^+Y^- \\ Y^-Y^+ \\ Y^-Y^- \end{cases}$$

15. The Pauli principle requires that, for half-integer spins, only antisymmetric states are allowed. With the four basis states one can build only one antisymmetric state, namely the so called "spin singlet":
$$\frac{1}{\sqrt{2}}\psi_{0,0,0,0}\otimes\left(Y^+Y^- - Y^-Y^+\right)$$

12.5. Exercises to Chapter 9.

1. Total energy of an H_{10}^{n} molecule.

In the lecture, we have computed the single-particle energy levels for an electron moving in a one-dimensional crystal with N atoms (lattice constant: a):

$$E_0^i = E_0 - 2\cdot A\cdot \cos(k^i\cdot a) \quad k^i = \frac{2\pi}{Na}\cdot i$$

E_0 is the value of the energy of the atomic-like state, $A > 0$. The index i labels the energy levels: it covers the range 1 to 10, hosting 10 possible eigenstates. The maximum number of electrons we can place in the crystal is 20, as we can occupy each level with two electrons (Pauli principle).

- 1. Compute the possible k^i-values for a H_{10}-molecule.
- 2. Compute the total energy of the neutral molecule H_{10}^{10}, using the shell model.
- 3. From the total energy, obtain the binding energy of the H_{10}^{10}- molecule.
- 4. Can you predict some magic filling numbers n for which the molecule H_{10}^{n} is particularly stable?

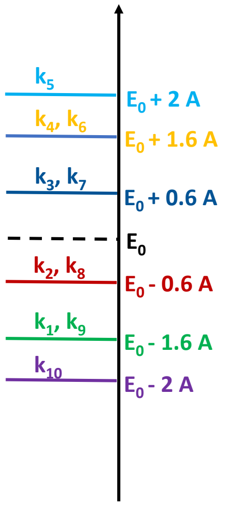
Solution.

- 1. The specification of the general formula to $N = 10$ writes

$$k^i = k^i = \frac{2\pi}{10\cdot a}\cdot i$$

The resulting energy levels are summarized in the following figure:

i	k^i		$E_i = E_0 - 2A\cos(k^i\,a)$
1	$\frac{2\pi}{10\,a}1 = \frac{\pi}{5a}$		$E_0 - 1.62A$
2	$\frac{2\pi}{5a}$		$E_0 - 0.62A$
3	$\frac{3\pi}{5a}$		$E_0 + 0.62A$
4	$\frac{4\pi}{5a}$		$E_0 + 1.62A$
5	$\frac{\pi}{a}$		$E_0 + 2A$
6	$\frac{6\pi}{5a}$		$E_0 + 1.62A$
7	$\frac{7\pi}{5a}$		$E_0 + 0.62A$
8	$\frac{8\pi}{5a}$		$E_0 - 0.62A$
9	$\frac{9\pi}{5a}$		$E_0 - 1.62A$
10	$\frac{2\pi}{a}$		$E_0 - 2A$



Left: the k^i values and the correspoding energy eigenvalues. Right: the energy eigenvalues marked along the vertical. Some k^i-values produces the same energy eigenvalue, which is, accordingly, more than once degenerate.

We find the following energy levels, with their degeneracy given in brackets

$$E_0 + 2A(1) \quad E_0 + 1.62A(2) \quad E_0 + 0.62A(2) \quad E_0 - 0.62A(2) \quad E_0 - 1.62A(2) \quad E_0 - 2A(1)$$

For computing the total energy of the neutral molecule H_{10}^{10}, fill the energy eigenvalues starting from the lowest one, and taking into account their degeneracy and the Pauli principle.

– 2 electrons in the level $E_0 - 2A$. These electrons contribute the energy

$$2 \cdot (E_0 - 2A) = 2E_0 - 4A$$

– 4 electrons in

$$E_0 - 1.62A \quad \rightarrow \quad 4E_0 - 6.48A$$

– 4 electrons in

$$E_0 - 0.62A \qquad \rightarrow \quad 4E_0 - 2.48A$$

The total enery of the 10-electrons molecule amounts to

$$E_{tot} = (2E_0 - 4A) + (4E_0 - 6.48A) + (4E_0 - 2.48A) = 10E_0 - 12.96A.$$

• 3. The binding energy is

$$E_{tot} - 10 \cdot E_0 = -12.96 \cdot A$$

• 4. From the lecture we know that we obtain a 'magic filling number' whenever one level is full. The level $E_0 - 2A$ is not degenerate, so we can fill 2 electrons there. The next magic filling number is the one for which the level $E_0 - 1.62A$ is full. This level is filled with 6 electrons in total (2 in $E_0 - 2A$ and four in $E_0 - 1.62A$). We obatain therefore the magic filing numbers

$$n_{\text{magic}} = 2, 6, 10, 14, 18$$

12.6. Exercises to Chapter 10.

1. One dimensional SE with Dirac-delta potential spike.

Find the transmission coefficient of a Dirac-delta potential barrier

$$\Phi(x) = \Phi_0 \cdot a \cdot \delta(x)$$

a is a characteristic length that mimics finite width of a realistic potential barrier. Φ_0 is a parameter that mimics the finite height of a realistic potential barrier.

Solution.

• The SE reads

$$\frac{\partial^2}{\partial x^2}\psi(x) + \frac{2m}{\hbar^2} \cdot (E - \Phi_0 \cdot a \cdot \delta(x)) \cdot \psi(x) = 0$$

This is an ordinary linear differential equation in one variable. One coefficient depends on the variable x, so that the solution is not straightforward. We notice that right and left from the singularity the potential is simple (in the specific case exactly vanishing) so that there we can solve the SE exactly. In the regions $|\, x \,| > 0$ the SE to be solved is

$$\frac{\partial^2}{\partial x^2}\psi_{L,R}(x) + \underbrace{\frac{2m}{\hbar^2} \cdot E}_{\doteq k^2}\psi_{L,R}(x) = 0$$

$\psi_{L,R}$ indicating the sought-for solutions for $x < 0$ (L) and $x > 0$ (R). The general solutions in the regions L and R read:

$$\psi_L(x) = A\frac{1}{\sqrt{2\pi}}e^{ikx} + B\frac{1}{\sqrt{2\pi}}e^{-ikx}$$

$$\psi_R(x) = C\frac{1}{\sqrt{2\pi}}e^{ikx} + D\frac{1}{\sqrt{2\pi}}e^{-ikx}$$

We use basis states which are, strictly speaking, non-square integrable and are normalized by the Dirac condition

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-ikx} \cdot \frac{1}{\sqrt{2\pi}}e^{ik'x} = \delta(k - k')$$

We hope that, despite the infinite norm, a procedure can be applied to obtain finite results. The meaning of these basis states is shown by explicitly computing the ”current” associated with them. As an example the Schrödinger current associated with the basis state $\frac{1}{\sqrt{2\pi}}e^{i\cdot k\cdot x}$ amounts to

$$\frac{\hbar}{2im}\Big(\overline{\frac{1}{\sqrt{2\pi}}e^{i\cdot k\cdot x}}\frac{\partial}{\partial x}\frac{1}{\sqrt{2\pi}}e^{i\cdot k\cdot x}-\frac{1}{\sqrt{2\pi}}e^{i\cdot k\cdot x}\frac{\partial}{\partial x}\overline{\frac{1}{\sqrt{2\pi}}e^{i\cdot k\cdot x}}\Big)=\frac{1}{2\pi}\cdot\frac{\hbar k}{m}$$

The plane wave basis state $\frac{1}{\sqrt{2\pi}}e^{i\cdot k\cdot x}$ produces a steady flow of probability in the direction specified by the sign of k.

As we want to describe a situation where a particle travel from the left toward the barrier, it is possibly transmitted and/or reflected, we allow non-vanishing coefficients A – describing the incident wave – B – describing the reflected wave and C – describing the transmitted wave – but we set $D=0$, i.e. we do not allow a source of particles in the R-segment. So far A,B,C are not specified and we need some physical conditions on how to determining A,B,C – or at least on how to relate them. One of this physical – boundary – conditions consists in requiring the continuity of the wave functions over the entire configuration space, even at singularities of the potential – the probability of a particle being at some site cannot depend on how one approaches the site. This means that

$$\psi_L(0)=\psi_R(0)$$

which leads to the first relation between A,B,C:

$$A+B=C$$

As delta functions can only be dealt with upon integrating over them, we take into account the singularity of the potential by integrating over an interval containing the origin and then letting the interval approach zero: (we call $\frac{2m}{\hbar^2}\Phi_0\equiv k_0^2$):

$$\underbrace{\int_{-\varepsilon}^{+\varepsilon}dx\cdot\psi''(x)}_{[\psi'_R(0)-\psi'_L(0)]}-\underbrace{\int_{-\varepsilon}^{+\varepsilon}dx\cdot k_0^2\cdot a\cdot\delta(x)\cdot\psi(x)}_{k_0^2\cdot a\cdot\psi_L(0)}+\underbrace{\int_{-\varepsilon}^{+\varepsilon}dxk^2\psi(x)}_{0}=0$$

From this equation one can read out the second boundary condition:

$$[\psi'_R(0)-\psi'_L(0)]-k_0^2\cdot a\cdot\psi_L(0)=0$$

which produces a further relation between A,B,C:

$$\begin{cases} A+B=C \\ ik\cdot[C-A+B]-k_0^2\cdot a(A+B)=0 \end{cases}$$

This is a linear system of equation for the sought for coefficients A,B,C and k (i.e. E). The system has solutions for any k: this means that the energy is not quantized. The energy spectrum builds a continuum of values. The solutions for A,B,C are

$$\frac{C}{A}=\frac{2ik}{2ik-k_0^2\cdot a}$$

$$\frac{B}{A}=\frac{k_0^2\cdot a}{2ik-k_0^2a}$$

Finally, by requiring that

$$|A^2|+|B|^2+|C|^2=1$$

the coefficient A can be determined up to a phase factor (we do not need explicitely this step to draw physically relevant conclusions, so we will avoid this algebra). We have therefore found the solution of the SE with a delta-like barrier:

$$\begin{aligned}\psi_L(x) &= \frac{A}{\sqrt{2\pi}}\cdot e^{ikx}+\frac{A}{\sqrt{2\pi}}\frac{k_0^2\cdot a}{2ik-k_0^2a}e^{-ikx}\\ \psi_R(x) &= \frac{A}{\sqrt{2\pi}}\frac{2ik}{2ik-k_0^2a}\cdot e^{ikx}\end{aligned}$$

The appearance of a finite coefficient C in the solution allows to compute a finite transmitted Schrödinger current:

$$J_T = \frac{|A|^2}{2\pi} \cdot \left| \frac{2ik}{2ik - k_0^2 \cdot a} \right|^2 \frac{\hbar k}{m}$$

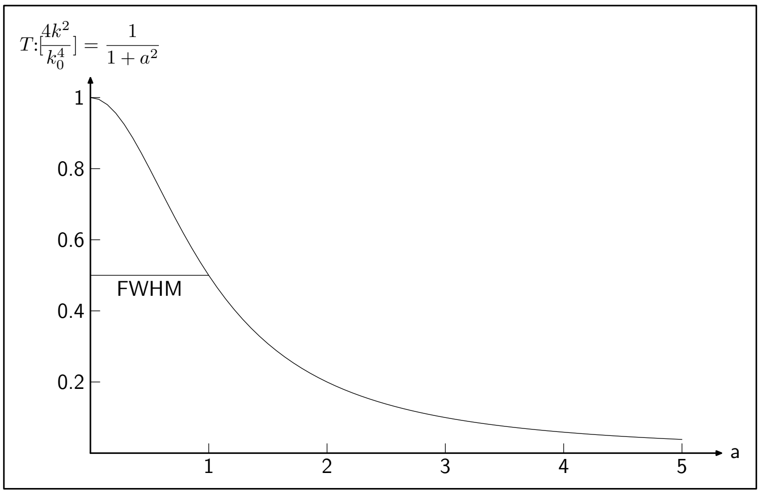
The transmission coefficient is defined as

$$T \doteq \frac{J_T}{J_I}$$

$J_I = \frac{|A|^2}{2\pi}\frac{\hbar k}{m}$ being the current arriving from the left hand side to the barrier. We find

$$T = \left| \frac{2ik}{2ik - k_0^2 \cdot a} \right|^2 = \frac{4k^2}{k_0^4} \cdot \frac{1}{a^2 + \frac{4k^2}{k_0^4}}$$

Notice that both incident and transmitted waves are non-square integrable, but when a ratio of currents is computed, a technically diverging normalization constant is divided out and the final result is free of unphysical divergencies. Our result about the finite transmission coefficient tells a very non-classical behaviour: we observe a finite probability for a particle crossing a classically forbidden spatial range. This phenomenon is known as quantum mechanical tunnelling. The transmission probability decay with a following a Lorentzian curve with full width at half maximum of $\frac{2k}{k_0^2}$. In term of the parameter E and Φ_0 the FWHM is $\sqrt{\frac{\hbar^2 E}{m \cdot \Phi_0^2}}$.



Transmission coefficient for a delta-like barrier. a is in units of $\frac{2k}{k_0^2}$.

For an electron ($m = 0.9 \cdot 10^{-30} kg$) with 1 eV energy, a barrier height of 4 eV produces a characteristic length of $\sqrt{\frac{\hbar^2 E}{m \cdot V_0^2}} \sim 0.6 nm$. At this length, the transmission coefficient is reduced by a factor of 2. Sizeable currents are still obtained in artificial lattices with barrier width in the nm range.

12.7. Exercises to Chapter 11.

1. Pauli equation: spin $\frac{1}{2}$ in magnetic field.

Let us assume that we have prepared a spin $\frac{1}{2}$ system in the state

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$$

at the time $t=0$.
- 1. Show that this state is an eigenstate of S_x.
- 2. Which value will a measurement of the z-component of the spin operator detect in this state?
- 3. Suppose now that the spin, obeying this initial condition, is subject to a uniform magnetic field pointing in the z-direction. Compute the spin polarization vector as a function of t.

Solution.

• 1.

$$\frac{\hbar}{2}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix}=\frac{\hbar}{2}\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix}$$

The given state is an eigenstate of S_x to the eigenvalue $\frac{\hbar}{2}$.
• 2. We write

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}=\frac{1}{\sqrt{2}}(Y^{+}+Y^{-})$$

This writewise makes clar that the sate is a superposition of a state witrh spin alonf $+z$ and a state with the spin along $-z$. A measurement of S_z will detect

$$\hbar\cdot\frac{\hbar}{2}\qquad-\hbar\cdot\frac{\hbar}{2}$$

with the same probability $\frac{1}{2}$.
• 3. The Pauli-Equation for $B_x=B_y=0$ is:

$$i\hbar\cdot\begin{pmatrix}\dot{C}_{+}\\\dot{C}_{-}\end{pmatrix}=\mu_B\cdot\begin{pmatrix}B_z & 0\\0 & -B_z\end{pmatrix}\cdot\begin{pmatrix}C_{+}\\C_{-}\end{pmatrix}$$

From this matrix equation we can obtain two ordinary differential equations:

$$i\hbar\dot{C}_{+}=\mu_B\cdot B\cdot C_{+}$$
$$i\hbar\dot{C}_{-}=-\mu_B\cdot B\cdot C_{-}$$

We solve the upper one:

$$i\hbar\dot{C}_{+}=\mu_B\cdot B\cdot C_{+}\Longleftrightarrow$$
$$\frac{\dot{C}_{+}}{C_{+}}=\frac{\mu_B\ B_z}{i\hbar}=-i\frac{\mu_B\ B_z}{\hbar}$$

So we can integrate both sides and obtain the solution of the differential equation:

$$C_{+}=a\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}\cdot t}$$

With the same calculation for the second equation we obtain:

$$C_{-}=b\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}\cdot t}$$

The general solution is the superposition

$$Y(t)=a\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{+}+b\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{-}$$

The integration constants are determined from the initial condition:

$$Y(t=0)=aY^{+}+bY^{-}=!\frac{1}{\sqrt{2}}(Y^{+}+Y^{-})$$

We read out

$$a=b=\frac{1}{\sqrt{2}}$$

and the special solution fulfilling the initial conditions writes

$$Y(t)=\frac{1}{\sqrt{2}}\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{+}+\frac{1}{\sqrt{2}}\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{-}$$

The spin polarization vector is defined as

$$\mathbf{P}=\left(<\sigma_x>_{Y(t)},<\sigma_y>_{Y(t)},<\sigma_z>_{Y(t)}\right)$$

We recall the rules of computation with Pauli matrices:

$$\sigma_x Y^{+}=Y^{-}\quad \sigma_x Y^{-}=Y^{+}\quad \sigma_y Y^{+}=i\;Y^{-}\quad \sigma_y Y^{-}=-i\;Y^{+}\quad \sigma_z Y^{+}=Y^{+}\quad \sigma_z Y^{-}=-Y^{-}$$

Than we compute:

$$\begin{aligned}
<\psi(t),\sigma_x\psi(t)> &= <\frac{1}{\sqrt{2}}\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{+}+\frac{1}{\sqrt{2}}\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{-},\frac{1}{\sqrt{2}}\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{-}+\frac{1}{\sqrt{2}}\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{+}> \\
&= <\frac{1}{\sqrt{2}}\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{+},\frac{1}{\sqrt{2}}\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{+}>+<\frac{1}{\sqrt{2}}\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{-},\frac{1}{\sqrt{2}}\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot Y^{-}> \\
&= \frac{1}{\sqrt{2}}\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}\cdot t}\cdot\frac{1}{\sqrt{2}}\cdot e^{+i\frac{\mu_B\cdot B}{\hbar}\cdot t}<Y^{+},Y^{+}>+\frac{1}{\sqrt{2}}\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}\cdot t}\,\frac{1}{\sqrt{2}}\cdot e^{-i\frac{\mu_B\cdot B}{\hbar}\cdot t}<Y^{-},Y^{-}> \\
&= \frac{1}{2}\cdot e^{+2i\frac{\mu_B\cdot B}{\hbar}\cdot t}+\frac{1}{2}\cdot e^{-2i\frac{\mu_B\cdot B}{\hbar}\cdot t}=\frac{1}{2}\left(e^{+2i\frac{\mu_B\cdot B}{\hbar}\cdot t}+e^{-2i\frac{\mu_B\cdot B}{\hbar}\cdot t}\right) \\
&= \cos\left(2\,\frac{\mu_B\cdot B}{\hbar}\cdot t\right)
\end{aligned}$$

In a similar calculation we obtain

$$P_y=<\psi(t),\sigma_y\psi(t)>=\sin\left(2\,\frac{\mu_B\cdot B}{\hbar}\cdot t\right)\qquad P_z=<\psi(t),\sigma_z\psi(t)>=0$$

Therefore the spin polarization vector is:

$$\mathbf{P}=\begin{pmatrix} P_x\\ P_y\\ P_z\end{pmatrix}=\begin{pmatrix}\cos\left(2\,\frac{\mu_B\cdot B}{\hbar}\cdot t\right)\\ \sin\left(2\,\frac{\mu_B\cdot B}{\hbar}\cdot t\right)\\ 0\end{pmatrix}.$$

2. Optical transistions from the $^2P_{1/2}$-level to the $^2S_{1/2}$-level

We consider the two basis states building the $^2S_{\frac{1}{2}}$ level as final states of an optical transition. The two initial states are the two states building the $^2P_{\frac{1}{2}}$ level.

- 1. Find the wavefunctions for the initial and final states from the corresponding tables.
- 2. The operator O_{light} of left circularly polarized light acts as following:

$$O_{light}Y_l^m=cY_{l\pm 1}^{m-1}$$

$$O_{light}Y_{\pm}=Y_{\pm}$$

(c is an unknown constant arising from the scalar product over the radial parts of the wave functions).
Compute all matrix elements for the possible transitions.

$$\left(\psi_{J_z=+\frac{1}{2}}^{S_{\frac{1}{2}}},[O_{light}]\psi_{J_z=+\frac{1}{2}}^{P_{\frac{1}{2}}}\right)$$

$$\left(\psi_{J_z=-\frac{1}{2}}^{S_{\frac{1}{2}}},[O_{light}]\psi_{J_z=+\frac{1}{2}}^{P_{\frac{1}{2}}}\right)$$

$$\left(\psi_{J_z=+\frac{1}{2}}^{S_{\frac{1}{2}}},[O_{light}]\psi_{J_z=-\frac{1}{2}}^{P_{\frac{1}{2}}}\right)$$

$$\left(\psi^{S_{\frac{1}{2}}}_{J_z=-\frac{1}{2}}, [O_{light}]\psi^{P_{\frac{1}{2}}}_{J_z=-\frac{1}{2}}\right)$$

Summarize the results in a energy-level-diagram.

- 3. Compute the spin polarization vector of those electrons that, after having made the transition, are in the level ${}^2S_{\frac{1}{2}}$. Hint: which basis state is populated, with which probability?
- 4. Assume now that the ${}^2P_{\frac{3}{2}}$-states also contribute to the transition. Compute the spin polarization vector of those electrons that have made the transitions

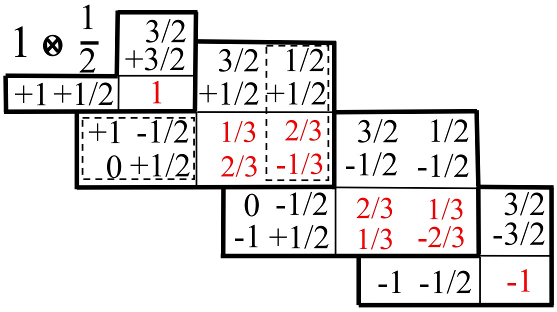
$$ {}^2P_{\frac{1}{2}} \rightarrow {}^2S_{\frac{1}{2}} \quad \text{or} \quad {}^2P_{\frac{3}{2}} \rightarrow {}^2S_{\frac{1}{2}} $$

Solution.

- 1. The ${}^2S_{\frac{1}{2}}$-level has two eigenstates

$$ |J=\frac{1}{2}, J_z=+\frac{1}{2}, L=0>=Y_0^0Y^+ \qquad |J=\frac{1}{2}, J_z=-\frac{1}{2}, L=0>=Y_0^0Y^- $$

From the Table



$1 \otimes \frac{1}{2}$	3/2 +3/2
+1 +1/2	1

	3/2 +1/2	1/2 +1/2
+1 -1/2	1/3	2/3
0 +1/2	2/3	-1/3

	3/2 -1/2	1/2 -1/2
0 -1/2	2/3	1/3
-1 +1/2	1/3	-2/3

	3/2 -3/2
-1 -1/2	-1

Table of the Clebsch-Gordan coefficients (in red) for $\mathbf{1}\oplus\frac{\mathbf{1}}{\mathbf{2}}$.

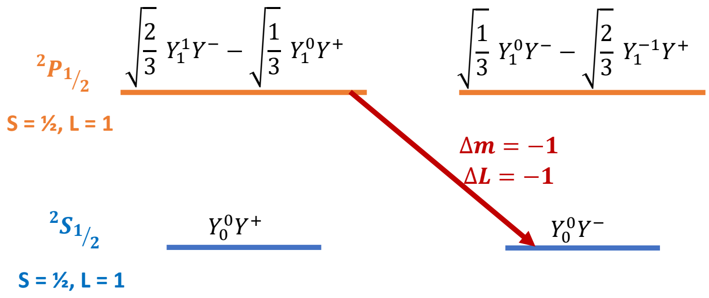
we obtain, for the two basis states of the ${}^2P_{\frac{1}{2}}$-level:

$$ |J=\frac{1}{2}, J_z=+\frac{1}{2}, L=1>=\sqrt{\frac{2}{3}}Y_1^1Y^- - \sqrt{\frac{1}{3}}Y_1^0Y^+ $$

$$ |J=\frac{1}{2}, J_z=+\frac{1}{2}, L=1>=\sqrt{\frac{1}{3}}Y_1^0Y^- - \sqrt{\frac{2}{3}}Y_1^{-1}Y^+ $$

Because of the rules of the application of the light operator, only one matrix element is different from zero. We compute it explicitely:

$$ \begin{aligned} <J=\frac{1}{2}, J_z=-\frac{1}{2}, L=0\,|\,[O_{Light}]\,|\,J=\frac{1}{2}, J_z=+\frac{1}{2}, L=1> &= <Y_0^0Y^-\,|\,c\sqrt{\frac{2}{3}}Y_0^0Y^-> \\ &= c\sqrt{\frac{2}{3}} \end{aligned} $$



The only transition possible with left-circularly polarized light is marked in the energy diagram.

- 3. there is only one final state: the spin polarization vector of the electrons excited into the $^2S_{\frac{1}{2}}$ level is

$$\mathbf{P} = (0, 0, -1)$$

- 4. The transition

$$^2P_{\frac{1}{2}} \rightarrow {}^2S_{\frac{1}{2}}$$

produces ”down” spins with probability

$$|c|^2 \cdot \frac{2}{3}$$

The transition

$$^2P_{\frac{3}{2}} \rightarrow {}^2S_{\frac{1}{2}}$$

produces spin ”up” electrons with probability

$$|c|^2 \cdot 1$$

and ”down” electrons with probability

$$|c|^2 \cdot \frac{1}{3}$$

The net spin polarization resulting from the combined trasitions is along the z-directions and writes (the unknown constant $|c|^2$ cancels out)

$$P_z = \frac{1 - \frac{1}{3} - \frac{2}{3}}{1 + \frac{1}{3} + \frac{2}{3}} = 0$$