Introduction to Quantum Informatics

WANTED

DEAD **&** ALIVE

Schrödinger's cat

$$\frac{1}{\sqrt{2}}(|\mathrm{DEAD}\rangle + |\mathrm{ALIVE}\rangle)$$

Levente Szabó

Introduction to Quantum Informatics

Pécs

2019

The Introduction to Quantum Informatics course material was developed under the project EFOP 3.4.3-16-2016-00005 "Innovative university in a modern city: open-minded, value-driven and inclusive approach in a 21st century higher education model".

EFOP 3.4.3-16-2016-00005 "Innovative university in a modern city: open-minded, value-driven and inclusive approach in a 21st century higher education model".

Levente Szabó

Introduction to Quantum Informatics

Pécs

2019

Az Introduction to Quantum Informatics tananyag az EFOP-3.4.3-16-2016-00005 azonosító számú,
„Korszerű egyetem a modern városban: Értékközpontúság, nyitottság és befogadó szemlélet egy 21. századi felsőoktatási modellben” című projekt keretében valósul meg.

EFOP-3.4.3-16-2016-00005 Korszerű egyetem a modern városban:
Értékközpontúság, nyitottság és befogadó szemlélet
egy 21. századi felsőoktatási modellben

Introduction to quantum information and quantum cryptography

Dr. Levente Szabó

WANTED

DEAD **&** ALIVE

Schrödinger's cat

$$\frac{1}{\sqrt{2}}(|\mathrm{DEAD}\rangle + |\mathrm{ALIVE}\rangle)$$

”*Everything we call real is made of things that cannot be regarded as real.”*
Niels Bohr

Contents

Chapter 1

Foreword

Quantum mechanics is the fundemental theory of all modern physics. As such it is a significant part of our understanding of Nature. It employs sophisticated mathematical models whose interpretation still poses unresolved physical and philosophical questions. Some of these lead to phenomena which are unusual and counterintuitive from an everyday perspective. In spite of this, quantum mechanics is a very successful theory. While its grounding fathers considered it as a theory for multipartite systems not verifyable on the level of individual physical systems, the formidable development of experimental technology (especially that of quantum optics) in the last decades has made the direct observation of these conterintuitive phenomena viable. And indeed: quantum mechanics appears to be valid for individual physical systems. Moreover, the accessibility of quantum phenomena seems to find its way to practical applications. The paradigm shift in physics introduced by quantum mechanics seems to be repeated in the field of information theory and information processing: quantum information now has a well-established reputation amongst future and emerging technologies. While quantum random generators and some quantum ciphers are commercial products already, yet there are still a lot of details to be better understood.

The goal of the present textbook is to provide a brief inside into some of the basic elements of quantum information to students taking part in a university training at the Faculty of Engineering and Information Technology at University of Pécs. Most of these attendees have never been trained in the field of physics especially in quantum physics, hence the overview of the fundamental quantum mechanics given in the introductory chapter is a bit wider than in other books with similar aims.

Quantum information science is often deeply related to the fundamental aspects of quantum mechanics. These aspects sometimes lead to paradoxical consequences when viewed from a classical physical perspective. Our other aim was to give a deeper understand of such paradoxes.

Chapter 2

Introduction to quantum mechanics

In this chapter we draw some of the basic elements of quantum physics. We will mostly follow *Prof. Dr. Tamás Geszti*'s guideline presented in his excellent university quantum physics book[1] (written in hungarian language). Actually this chapter includes the translation of some chapters from Tamás Geszti's book on quantum mechanics. We do not know any person who can give a better summary of the knowledge we need for understanding the fundamentals of our subject, than Prof. Geszti. Hence instead of trying to compete with his work, we just try to translate its relating parts from hungarian to english.

2.1 The beginning of the quantum theory

2.1.1 Planck's radiation law

By the end of the 19th century the building of the classical physics had been complete. Faraday and Maxwell grounded the discipline of eletrodynamics including the theory of emission and propagation of electromagnetic waves. Later this theory was proved by Hertz's experiments. Soon after it became evident that both light and thermal radiation belong to the family of this kind of waves. Simultaneously thermodynamics and statistical physics were born to start to discover the microscopical structure of the matter. Some dramatic realizations of Dalton and Avogadro denoted also this direction.

[1] titled *Kvantummechanika*

Coupling of thermal radiation and thermodynamics led to an incredibly seminal synthesis started by the theoretical investigation of Kirchoff. By that time from Bunsen's and others' experiments spectral structures of the thermal radiation had been known for many kinds of material or in other words the energy current $e(\nu,T)\mathrm{d}\nu$ emitted into a frequency interval $\mathrm{d}\nu$ around a given frequency ν by some material held at a temperature T was known. From this energy current, emission spectra was defined for different kinds of material. It was also known that from the incoming radiation, materials absorb the energy with a speed $a(\nu,T)u(\nu,T)\mathrm{d}\nu$ which is proportional to the energy density $u(\nu,T)\mathrm{d}\nu$ for each frequency range $\mathrm{d}\nu$, where the absorption spectrum $a(\nu,T)$ differs for each kind of material. $e(\nu,T)$ and $a(\nu,T)$ did not seem to be independent on each other: the colour (frequency) in which a material emits radiation with a high intensity at a given temperature T is the same colour which is absorbed by the material with a similar rate.

Kirchhoff was the person who realized this cannot be a coincidence, but this is the necessary condition for the thermal equilibrium between the thermal radiation and the emitter-absorber bodies with a temperature T in a cavity with the same temperature: if the rate between emission and absorption was not the same for each kind of material the energy would permanently flow from the strongly emitting materials towards the strongly absorbing ones which can be a stationary state but cannot be a thermal equilibrium. Conclusion:

$$\frac{e(\nu,T)}{a(\nu,T)} = u(\nu,T) \tag{2.1}$$

is a universal function or in other words $u(\nu,T)$ is the same for every material in the universe and has the following meaning: $u(\nu,T)\mathrm{d}\nu$ is the density of the radiating energy in the frequency interval $\mathrm{d}\nu$ in a cavity with a temperature T. Intensive experimental work was started to define this function. This work was founded on the observation according to which a very little hole on the wall of a cavity hardly disturbs the thermal equilibrium between the radiation and the walls of the cavity, hence a spectroscope directed onto the hole will show the spectral composition of the equilibratory radiation in the cavity. After several reflections the light incoming into the cavity through the hole, will be absorbed by the walls, sooner than it can get out of the cavity. This is the reason why this thermal radiation source is called *absolute black body*: a hole absorbing everything and emitting only radiation of thermal equilibrium, nothing else.

The spectral distribution of the outcoming energy density can be found out from the warming of a bolometer taken all along the screen of the spectroscope. Lots

of similar examinations were performed by many of the laboratories all over the world on different ranges of the spectrum and by the beginning of 20th century a $u(\nu, T)$ spectral function had been produced covering a wide frequency and temperature interval.

There was also a theoretical prediction according to which in a cavity (whose extension is much more greater than wavelengths that can be occured) with a volume V, in the $\mathrm{d}\nu$ interval, the number of the independent modes of the electromagnetic radiation with a propagation speed c^2 equals $V(8\pi/c^3)\nu^2\mathrm{d}\nu$. Each of these modes is considered as a harmonic oscillator having an energy of k_BT according to the *equipartition theorem*, where $k_B = 1.38 \times 10^{-23}J/K$ is the Boltzmann constant. From this it follows that the energy density belonging to each unit of volume $u(\nu, T) = (8\pi/c^3)k_B\nu^2T$.[3] This formula had to be altered. From a theoretical viewpoint, integrating this expression on the range of 0 to ∞, we get a radiational energy with an infinite value which is impossible. (The appearance of the infinity comes from the contribution of high frequencies, hence it is called "ultraviolett catastrophe".) The solution of the problem can be read out from the experimental results: beyond a certain frequency $\nu_{max} = BT^4$, function $u(\nu, T)$ is cut down. Hence the radiational energy integrated according to the frequency is finite: $U(T) \approx V(AB/3)T^4$. This result is in line with Stefan-Boltzmann law according to which T^4 dependence is a consequence of the general properties of pressure of light.

The point of the story was written by Max Planck who found out the physical meaning of the factor B and quantum theory was born. According to Planck's explanation every harmonic oscillator (including the electromagnetic radiational modes) can absorb energy in $h\nu$ *quanta* proportional to the frequency, where

$$h = 6.6 \times 10^{-34}Js \tag{2.2}$$

is the *Planck constant*. Evidently there are modes whose frequency is too big to take the energy belonging to the given mode, because the value of k_BT (according to the equipartition theorem) does not reach the value of the energy quantum $h\nu$. So at a given temperature T, only those wave modes (with a quasi continuous distribution) take part in the thermal excitation where $h\nu < h\nu_{max} = k_BT$. From this it follows the factor of the place of the frequency cuts $B \approx k_B/h$. Now we can replace the Rayleigh-Jeans law with the precise fromula called *Planck's law*:

[2]Where c denotes the speed of light

[3]This is the *Rayleigh-Jeans law*.

[4]*Wien's displacement law* because the location of the maximum is displaced proportional to the T increase

$$u(\nu) = \frac{8\pi}{c^3} \frac{h\nu^3}{e^{\frac{h\nu}{k_B T}} - 1}. \tag{2.3}$$

2.1.2 The photoelectric phenomenon: Lénárd and Einstein

The next appearance of the Planck constant h was related to the experiment of the Hungarian scientist Fülöp Lénárd who got a Nobel Prize for it. The relating formula was discovered by Einstein. The essence of the photoelectric phenomenon is that electrons can be disengaged from a metal electrode by light. In his experiment Lénárd measured the energy of the disengaged electrons (denoted by $E_{photoel}$) and realized it did not depend on the intensity of the light but depended on the *colour* of the light only.

Expressing the colour via the frequency of the light, Einstein (1905) wrote the formula into the following shape:

$$E_{photoel} = h\nu - W, \tag{2.4}$$

where h is the well known Planck constant and W is the *exit work* which is a feature of the illuminated metal electrode. According to Einsein's explanation, which was not believed by Fülöp Lénárd till the end of his life but was a reference base for Eintein's Nobel Prize, the incident light consists of energy quanta or in other word "*photons*". In case of electron-emission each electron takes the energy of exactly one piece of photon. Some part of this energy covers the exit work W and the other part changes the kinetic energy of the emitted electron: $E_{photoel} = h\nu - W$. Though photon hypothesis was believed by nobody for 20 years, after the borning of quantum mechanics Einstein turned out to be right.

2.1.3 The spectrum of atomic gases: from Rutherford's model to the Bohr model

In his experiment (1911), Rutherford analyzed the bounce of high energy positively charged α-particles of electrically neutral materials. It was found that the biggest part of the mass of any kind of material is crowded in the positively charged tiny nuclei surrounded by negatively charged easy electrons. According to Rutherford's explanation, the reason why an electron does not fall into the nucleus is that in every atom negatively charged electrons orbit around the positively

charged nucleus - as planets do around their sun in a real solar system. This is the *Rutherford model*, where centrifugal force keeps ballance with the Coulomb attraction. Unfortunately, this model has a mistake, namely the nucleus and the orbiting electron together form a rotating (viewed sideward: vibrating) electric dipole which system soon enough loses its energy by radiation, hence the electron should fall into the nucleus along a spiral shaped trajectory. The problem was solved by the analysis of the interaction between electrons and light (as we saw this approach was useful when photoelectric phenomenon was analyized). In the relating experiments the spectrum of the simpliest material was observed, namely: the monatomic hydrogen gas.

According to the measurements spectrum of the atomic gases consists of sharp lines belonging to a lot of frequency values. This chaos can be reduced by writing these frequencies as differences of two members:

$$\nu_{mn} = A_m - A_n. \tag{2.5}$$

Thinking about this, Niels Bohr was the person who found out (in 1913) that there can be a connection between equation 2.5 and Planck's quantum hypothesis or rather Einstein's photon hypothesis: let us multiply both sides of equation 2.5 by the Planck constant h and introduce the informative notation $hA_n = E_n$:

$$h\nu_{mn} = E_m - E_n, \tag{2.6}$$

whose obvious interpretation is that - from some unknown reason - certain sharply defined "stationary" electron trajectories - with an energy E_m, E_n - of the atom avoid the trap of the continuous dipole radiation[5]. However, once in a while, the atom emits or absorbs a photon with an energy $h\nu_{mn}$ by leaps. At that moment when it happens, the electron jumps from one of its stationary trajectories to another one. If we want to know what will exactly happen (as we know it can be either emission or absorbtion), we have to find out the relation between the energy of the initial trajectory and the energy of the final trajectory. Equation 2.6 expresses the law of the energy conservation in the process of the quantum leap.

Though everything we dealt with so far is valid for all kind of monatomic gases, in the case of Hydrogen there is more to be learn. According to Rydberg's empirical formula

[5]The existence of the stable electron trajectories with discreet (or in other word noncontinuous) energy values was proved by the *Frank-Hertz experiment* (in 1914).

$$E_n = -Ry/n^2, \tag{2.7}$$

where Ry is the Rydberg constant. In Bohr's interpretation this formula can be deduced if the electron orbits along a circle shaped trajectory and the radius r of the circle shaped stationary trajectories (which avoid the trap of the continuous radiation) are selected from the many of the trajectories allowed by the classical mechanics (but also containing radiation, thus condemned to collapse by eletrodynamics) by the following "quantum condition" (which refers to the angular momentum $L = pr = M_e vr$):

$$L = n\frac{h}{2\pi} \equiv n\hbar. \tag{2.8}$$

Here we introduced a (in quantum mechanics) commonly used notation: $\hbar = h/2\pi$. Its value $\hbar = 1.03 \times 10^{-34} Js$. From the calculation for the radius of the trajectory with a value $n = 1$ (***Bohr radius***) we get the following value:

$$r_B = \frac{4\pi\varepsilon_0\hbar^2}{M_e e^2} = 5.3 \times 10^{-11} m, \tag{2.9}$$

where $M_e = 0.91 \times 10^{-30} kg$ is the mass of the electron, $e = 1.6 \times 10^{-19} C$ is the charge of the electron and ε_0 is the dielectric permeability of the vacuum, its value in SI units is $8.85 \times 10^{-12} C^2/Jm$. In the end, for the factor of the equation 2.7, we get the following value:

$$Ry = \frac{e^4 M_e}{(4\pi\varepsilon_0)^2 2\hbar^2} \approx 2.18 \times 10^{-18} J \approx 13.6 eV. \tag{2.10}$$

Replacing the factors of the right side of equation 2.10 with the experimentally measured values we get the frequencies of the spectrum lines of monatomic hydrogen with a very high accuracy. According to the conventions of physics, an accordance with such a high accuracy is a sure sign for Bohr's quantum condition that it takes some important feature of the physical reality. Inserting 2π into the formula 2.8 may seem to be arbitrary, still it gave into Bohr's hand the key of this numeric accordance (just like a code to open a safe). The only thing we have to do is to try to understand how and why.

2.1.4 De Broglie: the interpretation of the Bohr formula with matter waves

As Einstein remarked, since according to Maxwell's electrodynamics a momentum E/c belongs to the energy E of an electromagnetic wave, also a photon with an energy $h\nu$ has a momentum $p = h\nu/c = h/\lambda$. From Einstein's point of view this was an essential step to consider photon as a full member of elementary particles.

In 1924, Louis de Broglie discovered a fantastic explanation for Bohr's quantum condition. His contexture was the following: Reversing Einstein's argument, he supposed that a motion of some kind of *matter wave* can be assigned for an electron moving with a momentum $p = M_e v$. If Einsten's connection

$$\lambda = \frac{h}{p} \tag{2.11}$$

holds true for the case of the mentioned matter wave, we get an amazingly simple explanation of Bohr's quantum condition. And indeed: altering equation 2.8, we can rewrite it into another shape:

$$2\pi r = n\frac{h}{p} = n\lambda, \tag{2.12}$$

where - as we know - n is an integer number, which means that the circulating wave closed into itself can continue its stationary undisturbed wave-motion perpetually. This selects Bohr's stationary orbits.

The 2.11 de Broglie relation, which creates a connection between the wavelength of the electron moving like a wave and the momentum of the electron moving like a flying ball is one of the fundamental connections in quantum mechanics. In addition to this it is true not only for photons or electrons, but for any of quantum motions too. Since momentum is not a scalar quantity, but a vector, we need to introduce the wave-vector $\vec{k}$, whoose length is $k = 2\pi/\lambda$ while its direction is normal to the wave front. Thus equation 2.11 can be softened into the following form:

$$\vec{p} = \hbar\vec{k}. \tag{2.13}$$

De Broglie found a wide variety of argument to make the existence of the matter

waves plausible. The most directed one is that if in case of some conditions a light wave acts like a particle (photon), then we must not be surprised by the behaviour of an electron when it acts like a wave in case of certain conditions.

There is a more sophisticated argument: if we know the energy of a vibrational quantum is related to the frequency of a vibration (as we saw in Planck's formula: $E = h\nu = \hbar\omega$), then we have to remember that in Einstein's relativity, energy E and the three components of momentum $\vec{p}$ together forms a four-vector and this is exactly the case when we consider the ensemble of pulsatance ω and the components of wave-vector $\vec{k}$. Thus if a temporal vibrational motion belongs to the energy, then a wave motion is related to the spacial momentum: the existence of the matter wave is required by the symmetry-world of relativity!

At first sight, this argument may seem to be extraordinary, but it is very effective: in modern physics it is a beloved toy to predict the existence of something built upon a symmetry and then to find the predicted object (or phenomenon, etc). The enormous heuristic power of relativity helped many times the advance of quantum theory in the past.

A lot of fundamental questions seem to be swept under the carpet:

- How can something move like a tiny ball, if it is a wave?
- Though we do not know what an electron is exactly, we are sure it is a charged object. In this case it should emit during a rotary motion. How does Bohr - de Broglie quantum condition ensure it does not emit the energy of its rotary motion instantly?
- What waves?

These important questions were solved in a surprisingly simple way at the end of the 1920's and quantum mechanics was born. The next section is devoted to clarify these problems including the experiments which proved the existence of the matter waves.

2.2 The elementary properties of the matter waves

2.2.1 Interference experiments from the electron to C_{60}

The most sightful wave phenomenon is the interference, where a wave can be divided to partwaves. These partwaves then go along different trajectories divided in space and/or in time conserving their vibrational phase which derives from their common origin. At the end of the proceed, after crossing they make each others more intens or weaker depending on the phase difference which was created during the separated fly.

The following tools are required in case we want to observe the interference:

- *a wave source:* if we deal with light we can use a lamp or a laser source, but if we want to analyse matter waves we have to use some hot matter which can be an electron cannon with a hot cathode and an accelerating electrode, or a stove vaporizing atoms and molecules, or we can use some kind of neutron source such as a reactor emitting neutrons or a spallation neutron source;
- *a beam splitter:* a screen with double slit, an optical lattice, or a semi-permeable mirror;
- *a space enough:* which is needed when we insert objects to influence the phase and amplitude of the part beams;
- by chance a *beam amplifier* which is just like a beam splitter;
- *a detector* or a system of detectors which observes the interference pattern.

The interference of light was first observed 200 years ago. For this, as a novelty, the existence of hard separable light-quanta was discovered in the 20th century. However, matter waves appear in the interference of elementary or complex matter-particles (like electrons, atoms or molecules) which was a shocking surprise at that time. The interference pattern can be observed even in the case of so weak beam intensity where it is evident that only one piece of particle propagates as a wave, divides on a beam splitter and interfere with itself, and the pattern can not be resulted by the collective motion of many particles.

For the first time, electron-interference was observed by Davisson and Germer. At that time, in 1927, this was a Nobel Prize experiment. Nowadays, it is routine accessory of electron microscopes.

Neutron interference became a routine technic in interference experiments in the seventies. We will give a deeply detailed discussion of this experiment later as an indroduction for the idea of quantumbits (or in short: qubits).

In the case of atoms or molecules, it is hard to create a source and to solve the detection too. If we deal with molecules an additional difficulty appears, namely the effect of the vibration and rotation of the molecule which smudges the interference. In 1999, the Zeilinger-Arndt experimental group in Vienna achieved sensational experiments in which interference of C_{60} and C_{70} fullerene-molecules and some other similar sized organic molecules was observed. The mass of the biggest molecules in these interference experiments was about $3 \times 10^{-24} kg$ which seems to be a hard breaking limit because matters consisting of heavier molecules can not be vaporized in a stove, hence there is no known way to create a wave source from them.

2.2.2 The superposition theorem

The electric field strength depending on space and time is the amplitude of the light-wave. The changing height of the surfice of the roughish water - relative to the level of the static water - is the amplitude of the water-wave. Conventionally, as Schrödinger introduced, the amplitude of the matter-wave is denoted by the symbol $\Psi(\vec{r}, t)$ and quantum mechanics is required to enlighten the physical content of this function.

In interference-phenomena, the amplitudes of the part waves are summed up. Being signed quantities (or - by chance - vectors), their resultant summation can both intensify and weaken depending on space and time. This effect results the typical interference fringes. Our detectors observe these phenomena via the *intensity* of the wave. In case of electrical or mechanical waves, the intensity is the absolute value square of the amplitude. Accepting - as a trial - this conception in the case of matter waves too, quantity $|\Psi(\vec{r}, t)|^2$ has to be considered the intensity of the wave. Using this, we can describe the interference of two part waves in the following way:

$$\begin{aligned} |\alpha\Psi_1(\vec{r}, t) + \beta\Psi_2(\vec{r}, t)|^2 = |\alpha|^2|\Psi_1(\vec{r}, t)|^2 + |\beta|^2|\Psi_2(\vec{r}, t)|^2 \\ +\alpha^*\beta\Psi_1^*(\vec{r}, t)\Psi_2(\vec{r}, t) + \alpha\beta^*\Psi_1(\vec{r}, t)\Psi_2^*(\vec{r}, t), \end{aligned} \tag{2.14}$$

where Ψ can be a complex number[6] and the operation of complex conjugation is

[6] It will be useful later.

denoted by the symbol $*$. Factors α and β in the *linear combination* $\alpha\Psi_1(\vec{r},t) + \beta\Psi_2(\vec{r},t)$ denote the possibility of the two part waves to take part in the interference with different weights. The first two members on the right side of equation 2.14 is the sum of the intensities of the two part waves weighted according to the linear combination. The interference itself is denoted by the last two members: the mixed products resulting a negative number where and when the two part waves vibrate in opposite phase (*destructive* interference) and resulting a positive number (*constructive* interference) if the two part waves vibrate in the same phase. Let us notice α and β are mostly complex, thus they can shift the interference pattern with their complex phase.

The following statement is experimentally verified for electromagnetic waves propagated in vacuum: Suppose we have two part waves evolving in time separately. After their time evolution, the sum of them is as exactly the same as a resultant wave in case we had the linear combination of them at the beginning of the time evolution. This is expressed mathematically by the linearity of Maxwell's equations. For water waves, linear combination can be applied just as a rough approximation. Thus it is extremly surprising to find it true in case of matter waves: the *superposition* of these waves is preserved during their time evolution with a huge accuracy. This experience is expressed by the *superposition theorem* which is one of the fundamental laws of quantum mechanics. Mathematically it means that an equation describing the propagation of the matter waves has to be linear. The 90-year-old superposition theorem has never been caught at fault in the micro-world, even in the most exotic circumstances.

On the other hand, macroscopic bodies never appear in a superposition of different places or states. If quantum mechanics was valid for us, it would be enough for us to take part in an interaction with a particle whose state is a superposition and our state would be also superposition. This absurdity appears in Schrödinger's famous "cat-metaphor": in a box, a cat-killer tool (controlled by a detector which clicks in case of observing a radiogenic α-particle) kills a cat. According to the linear quantum mechanics, while the α-particle is in the superposition of the states "already emitted" and "not emitted yet", due to the detector(cat killer tool)-cat interaction, the cat is in the superposition of the states "dead" and "alive" too.

No one has seen anything like this. The conclusion is obvious: for cats, people, macroscopic bodies laws of the linear quantum mechanics are not valid. We do not know yet if there is a continuous transit between the macro- and microworld. Nevertheless, it is very important to know that each known physical system is on one of the two coasts of this micro-macro abyss and never inside of it. Quantum mechanics describes the interactions between the two world (macro and micro)

with perfectly working rules. Making a measurement on a micro system with macroscopic detectors is a typical one of these interactions.

2.2.3 Wave packet, group velocity

The wave motion can seem to be the motion of a flying particle (just like the feather of a water wave). With the help of the superposition theorem, we can say that a *wave packet* takes shape from the superposition of waves with different wavelengths. This analogy is precise and significant. In a randomly created superposition, waves whose wavelengths are different extinguish each other in most of the places. Wave packets can take shape in places only where phase of the part waves coincide with each other. Since phase-velocities of the part waves are not the same, the place of the wave packet is moving. The velocity of this motion is called *group velocity*.

Its mathematics is very simple. Consider a plane wave propagating along the x-axes of a coordinate system. The wave number of this wave is $k = 2\pi/\lambda$ and it can be described mathematically in the following way:

$$\Psi_k(x,t) = \sin(\psi_k(x,t)) = \sin(kx - \omega t), \tag{2.15}$$

from where it can be seen that places with the same phase ψ propagate with a velocity $\mathrm{v}_{\mathrm{phase}} = \omega/k$. Let us assemble (create a superposition) different plane waves whose wave number k is a continuous variable and whose frequency depends on the wave number in the following way: $\omega = \omega(k)$. The resutlant wave amplitude can be calculated from a Fourier integral in which the components are integrated with some kind of $c(k)$ real weights:

$$\Psi_k(x,t) = \int c(k)\sin[kx - \omega(k)\,t]\,\mathrm{d}k. \tag{2.16}$$

Here comes the mathematical point. As we have already mentioned, the integrated sum of the components with different wavelength equals 0 in most of the x places and t times. Exceptions are the cohesive (x,t) pairs, where and when the k-derivative phase disappears, or in other words *there is no* oscillation:

$$\frac{\partial}{\partial k}[kx - \omega(k)\,t] = x - \frac{\partial\omega(k)}{\partial k}t = 0. \tag{2.17}$$

From this formula, it follows that these time-dependent places are moving with a velocity

$$\mathrm{v}_{\mathrm{group}} = \frac{\partial\omega(k)}{\partial k}, \tag{2.18}$$

which is called *group velocity*.

So far, we dealt with waves propagating along the x-axis only. However, in a real life situation a plane wave can be propagated along any of the directions denoted by a wave vector $\vec{k}$. In this case the group velocity can be calculated according to the following expression:

$$\vec{\mathrm{v}}_{\mathrm{group}} = \frac{\partial\omega(\vec{k})}{\partial\vec{k}}. \tag{2.19}$$

Only one step we need to do so as to believe that the motion of a wave packet has something to do with the motion of a pointlike particle (whose motion can be described by the tools of classial mechanics). Let us recall the Bohr - de Broglie rules of the wave-particle correspondence: $\hbar\omega = E$, $\hbar\vec{k} = \vec{p}$. Applying these rules, we can reshape equation 2.19:

$$\vec{\mathrm{v}}_{\mathrm{group}} = \frac{\partial E(\vec{p},\vec{r}\,)}{\partial\vec{p}} = \frac{\partial}{\partial\vec{p}}\left(\frac{p^2}{2m} + V(\vec{r}\,)\right) = \frac{\vec{p}}{m}, \tag{2.20}$$

which equals the velocity come from the classical mechanics!

As anyone can see here, the thing is really that how our world is working. Nevertheless the wave packet- picture is limited, because in the microworld wave packets are short-lived objects, they arise and feather. In certain exceptional circumstances (for example in case of harmonic vibrational motion), there can be wave packets which swing stably instead of feathering. This kind of wave packets are called coherent states.

2.2.4 Motion in a force field

We expect quantum mechanics to desribe also those situations which are well featured by classical mechanics, namely quantum mechanics has to be able to give a precise description of motions in a force field featured by a potential $V(\vec{r}\,)$.

From classical mechanics we know that in a case like this, there is an acceleration proprotional to the force $-\vec{\nabla}V$, while energy $E = p^2/2m + V(\vec{r})$ is a constant. If a body goes to a place where the potential $V(\vec{r})$ reaches its total energy E, or in other words it falls foul of the potential wall, it will bounces.

Though there is energy conservation in the world of waves too, the motion shows a wider variety of forms than classically. First of all, correspondently equation 2.11, where $V(\vec{r})$ changes (while energy E remains unchanged), wavelength changes too. Expressing it via wave number, we get that

$$k(\vec{r}) = \frac{1}{\hbar}\sqrt{2m(E - V(\vec{r}))}. \tag{2.21}$$

So, in case of moving along the force (potential energy decrease) the wave number of a localized wave packet increase. Or in other words, also its momentum increase, in line with the Newtonian mechanics. If we consider a nonlocalized wave packet which can be a self-contained wave feathering along a closed trajectory, we meet a nice generalization of Bohr's quantum condition: in this case a standing wave corresponds to the circulating motion along a q coordinate, because the length of the total trajectory is the integral multiple of the wavelength $\lambda(q)$. A fraction of $\mathrm{d}q/\lambda(q)$ of the wavelength can fit on a line segment $\mathrm{d}q$ at a given position q. So if $\oint \mathrm{d}q/\lambda(q) = n$ is an integer number, the motion will be stationary, because this is the condition of a stationary motion. Or in another shape:

$$\oint p(q)\mathrm{d}q = n\,h. \tag{2.22}$$

This expression was discovered by Sommerfeld before de Broglie's wave picture, hence it is called Bohr-Sommerfeld quantum condition.

In case of rebounding from a potential wall there is a shocking phenomenon, namely waves do not act like flying balls in classical mechanics. This is already known from classical optics: in case of total reflection, light gets into the reflective medium as an *evanescent* wave. This is exactly the case when we deal with matter wave: where $V(\vec{r}) > E$, kinetic energy $\hbar k^2/2m$ is negative. For waves, this means that the wave number k is *imaginary*, hence it can be written in this way: $k = i\kappa$, where κ is a real number. In this case, the wave-amplitude depends on the depth z of the influx according to the following formula: $e^{ikz} = e^{-\kappa z}$, or in other words, the amplitude decreases exponentially, instead of oscillating. What is it all about? If the "forbidden" - $V > E$ - interval is thin enough, then there is a little fraction of the matter wave which gets over the potential wall. This is called

tunnelling effect and has a lot of manifestations: this is the way for an α-particle to get out of the radioactive nucleus, and this makes it possible for the electric current to get through the oxidized surfaces of the light switch. In biological mutations and important enzyme reaction, the tunnelling of protons has a central role: without it, we would not have come into being.

2.2.5 Semiclassical motion and transition into the classical mechanics

Wave packets are extended and complex in general, thus we need to know its shape in detail to predict its behaviour. Nevertheles, in a limiting case, when the size of the wave packet is much smaller than the range of the motion (it means that the wave packet consists of short waves), motion of the wave packet can be derived from the equations of the classical mechanics. This is said to be the *semiclassical* approximation of the quantum mechanics. Important systems belonging to this limiting case are said to be *mesoscopic*, or in other words, they are on the limit of the microscopic and macroscopic range (between the two "world"). These systems are well testable experimentally. and they can be electrons moving in semiconductor structures (on a very low temperature). It is important for these kind of structures to be made of very clean materials.

At the very beginning of the quantum mechanics, connection between quantum and classical mechanics seemed to be understood in this way: in a certain limiting case (for instance in case of short wavelength) quantum mechanics turns into classical mechanics. This was called correspondence principle by Bohr. Formally, this principle works but our faith (in its content simplicity) has disappeared by this time. While wave packets show wave properties, namely it can interfere, its phenomena have to be described by quantum mechanics. Though the motion of a wave packet (in the mentioned limiting case) obey the law of classical mechanics, existence of classical mechanics is not near understood from the viewpoint of the quantum theory.

We have to step forward to understand the really classical physics. In the last quarter of the 20th century Zeh and Zurek reached a significant progress in understanding how coherence of the matter waves can vanishes while heading towards macroworld from microworld. The key element is the quantum mechanical description of the interaction with the environment. Though this result was great breakthrough, there are still open questions in this topic, especially in the field of quantum measurement. Clarification of these problems lays upon the new millenium.

2.2.6 Uncertainty relation

A wave packet is similar to a flying particle but not exactly the same thing. While a particle has definite position and momentum, in case of a wave packet, these quantities can be defined with a limited accuracy. Indeed, wave packets have a Δx width which can be considered the uncertainty of the location (for the sake of simplicity we restrict the motion to one dimension again). On the other hand, according to the equation 2.16, the wave packet consits of plane waves selected from a wave number interval whose width is Δk. From this it follows that wave packets have also a momentum uncertainty $\Delta p = \hbar \Delta k$.

These uncertainties are not independent of each other, because they are roughly inversely proportional to each other. Generalized mathematical details will be discussed later, but the content of this topic is fairly simple: if an averagely inert wave packet is made of plane waves with wave numbers from the interval 0 to $kmax \equiv \Delta k$, then the size of the wave packet is roughly the half of the shortest wavelength, that is

$$\Delta x \approx \lambda_{min}/2 = \pi/k_{max}. \tag{2.23}$$

Similarly, the scatter of the momentum is

$$\Delta p \approx \hbar k_{max}/2. \tag{2.24}$$

Their product is

$$\Delta x \Delta p \approx \hbar \pi/2. \tag{2.25}$$

The precise result derived in a future chapter yields an exact limit:

$$\Delta x \Delta p \geq \hbar/2. \tag{2.26}$$

Expression 2.26 is *Heisenberg's uncertainty relation* which was said to be the most important result of quantumm mechanics for a long time. By now we got used to it and use it routinely with many of its analougous formulae. One of them is the connection between the uncertainty of time and uncertainty of the energy: if we want to define the date of an event with an accuracy Δt, we have to assemble periodical vibrations from a frequency interval $\Delta\omega \geq 1/(2\Delta t)$ so as to get an

impulse with a length Δt. According to quantum mechanics $E = \hbar\omega$ of energy belongs to a vibration whose frequency is ω. Hence, for this case uncertainty relation is

$$\Delta t \Delta E \geq \hbar/2, \tag{2.27}$$

which shows that the energy of a particle (decayed in a finite time) can not be totally definite.

Let us mark: bounds ruled by the uncertainty relations would dessapeared in the $\hbar \to 0$ limiting case, which is one of the expressions of the correspondence theorem.

2.2.7 The size and energy of the ground state

Why are atoms so big? For the nucleus end elecrons, much less place would be enough. In addition to this, even the potential energy would be lower, because of the Coulomb attraction. However, in this case their kinetic energy would increase, because of the uncertainty relation. There is a most stable size, where the sum of the potential and kinetic energy is the smallest that can be. This is the quantum mechanical *ground state*.

More precisely, according to the uncertainty relation, the quadratic average deviation (from 0) of momentum increases to $\hbar^2/(2\,\Delta x)^2$ when the size is reduced to Δx. The value of the kinetic energy belonging to this size equals $\hbar^2/8m\,(\Delta x)^2$. The stable size and its ground state energy can be defined as the minimum of the total energy

$$E(\Delta x) \approx V(\Delta x) + \frac{\hbar^2}{8m\,(\Delta x)^2}. \tag{2.28}$$

The resulted value depends on the potential $V(x)$ which holds the system together. Let us try out this method for some cases. In case of a harmonic oscillator $V(x) = (m/2)\omega^2 x^2$. From this expression, it follows that the size belonging to the minimal energy is $\Delta x = \sqrt{\hbar/2m\omega}$ and the ground state energy ("zero point vibration") equals $E_{min} = \hbar\omega/2$. This result is too good to be true, but it is: if we solved the Schrödinger equation, we would see that our result obtained by a rough estimate is correct.

As another trial, let us consider the hydrogen atom. In Bohr's model, the potential of the attraction between proton and electron is $V(r) = -e^2/4\pi\varepsilon_0\, r$. Replacing Δx with r in the expression of the energy, for the place of the minimum we get a value of $r = \hbar 2\pi\varepsilon_0/e2m$, and for the energy $-e^4m/8\pi^2\varepsilon_0^2\hbar^2$ is the calculated value. This result is four times greater than the ground state energy of Bohr's model, but as for the order of megnitude, it is a correct result.

The strategy when we minimize an approximate energy formula (according to its free parameters, here: Δx and r) and get an estimate for the ground state energy is called *variational principle*. In the last 90 years, a significant part of our physical knowledge has been discovered by applying more and more sophisticated variational princiles. The presented primitive version can be named as bagger's variational principle.

2.2.8 Randomness and Born's statistical interpretation

There is one more important step we have to take: what does the intensity of waves mean? The direct answer: intensity is a thing measured by detectors, photography or observations of a fluorescent screen in interference experiments.

A more sophisticated answer turns out in those cases, when the intensity of the incoming matter wave is so small that maximum one piece of particle can be contained by the experimental setup or less. In a situation like this, experiments show a dramatical picture: one detector clicks only, one place on the screen flashes only, one point of the film's becomes black only at the same time and if the particle survives the detection, after the observation the whole matter wave will start to pass on, from the point where the particle was detected. The position of this point depends on the randomness. However from the random events, interference pattern stands out in time! This phenomenon is expressed mathematically by Born's rule:

$$|\Psi(\vec{r},\, t)|^2 d^3r \tag{2.29}$$

is the probability for a detector to find a particle – which corresponds to the matter wave with an amplitude $\Psi(\vec{r},\, t)$ – around a place denoted by $\vec{r}$, in a volume d^3r. In other words: $|\Psi(\vec{r},\, t)|^2$ is the *finding probability density* of the particle.

It is an independent experience (from things dealt with so far, it does not follow) that a particle can be found on one place only (here *or* there etc.) but never on

several places at the same time. These events mutually exclude each other. Hence the probability of finding the particle certainly *somewhere* in the space is the sum (or integral) of probabilities 2.29:

$$\int |\Psi(\vec{r},\, t)|^2 d^3r = 1. \tag{2.30}$$

This significant feature is called the *normalization* of the wave function.

At length, we can get used to the thing that the invisible tiny electron is not a yellow pillet, but it is a wave. Stranger properties of quantum mechanics starts from the randomness of the detectation. Experimentally the only possibility for us to come at matter waves is our detectors. And they yield random signs. This feature belongs to the wave function inseparately. Thus wave function is named *probability amplitude* too.

The really mysterious thing is that one detector fires only, every time! How does a detector know another one fires, hence it has to be quiet? This was first noticed by Einstein at a conference and in a few years topic of Einstein-Podolsky-Rosen correlations emerged from this remark[7]. Born's rule holds true with a very high accuracy in all cases where it can be checked (just as the superposition theorem).

In the formula 2.29, we can find the square of the absolute value of a *complex* number. If we multiply the wave function by a complex number $e^{i\chi}$ whose absolut value is a unit, the square of the absolut value of the wave function remains unchanged. This feature holds true for every combination which can be observed (measured), hence wave functions $\Phi(x)$ and $\Phi(x)e^{i\chi}$ are equivalent physically.

2.2.9 Standing waves: matter conservation and complex amplitude

Superposition theorem creates a nice trap situation. How should we describe a matter wave emerging at that moment when an electron is just being reflected from a reflective surface and it is present in the incoming and in the reflected wave at the sam time? Let us sum a plane wave heading to the right and another one heading to the left: the resultant wave is:

$$\sin(kx - \omega\, t) + \sin(-kx - \omega\, t) = 2\sin(\omega\, t)\cos(kx). \tag{2.31}$$

[7]This topic will be discussed later in the field of quantum entanglement.

This expression describes a standing wave disappearing twice per period. Question: what is with the law of matter conservation in such case? Where is the electron when the wave disappears? Rules discussed so far can not answer this queation correctly. Its solution was discovered by Schrödinger. He suggested to rewrite the wave amplitude via a *complex* function. In case of waves, vibrations, this is a commonly used mathematical trick. We can sum comlex numbers more easily than trigonometric functions, but in such cases one keeps in mind that the real part of the complex function is the reality. However, in this situation, our trick becomes reality: a remaining particle (as matter wave) has to be described really with a complex function and physical content is included by the whole function, not only its real part.

How does it work? The amplitude of a plane wave heading to the right is $e^{ikx-i\omega\,t}$ now. The superposition of the plane wave heading to the left and another one heading to the right is:

$$e^{ikx-i\omega\,t}+e^{-ikx-i\omega\,t}=2e^{i\omega\,t}\cos(kx). \tag{2.32}$$

It never disappears, it just rotates on the complex plane. In case of functions with a shape like this, the absolute value square, the wave intensity, or in other words the finding probability density does not change in time, because $|e^{-i\omega\,t}|^2=1$, so there is no problem with matter conservation anymore.

This is the key for us to solve the "no radiation" paradox which seemed to be unsolvable in Bohr's atom model: stationary electron orbits are actually standing waves with an amplitude of $e^{-i\omega\,t}\Phi(\vec{r})$. In these standing waves, distribution of the electric charge density remains unchanged in time, thus electromagnetic waves are not emitted.

However during a transition between two stationary orbits, the state of an electron (in the "language of waves") is the following superposition: $c_1e^{-i\omega_1\,t}\Phi_1(\vec{r})+c_2e^{-i\omega_2\,t}\Phi_2(\vec{r})$. Meanwhile the finding probabilty density of the electron – hence also the distribution of the electric charge density – changes in time in this way:

$$\begin{aligned}|c_1e^{-i\omega_1\,t}\Phi_1(\vec{r})+c_2e^{-i\omega_2\,t}\Phi_2(\vec{r})|^2=|c_1|^2|\Phi_1(\vec{r})|^2+|c_2|^2|\Phi_2(\vec{r})|^2+\\ 2\Re[c_1c_2^*\Phi_1(\vec{r})\Phi_2^*(\vec{r})]\cos(\omega_{12}\,t)+2\Im[c_1c_2^*\Phi_1(\vec{r})\Phi_2^*(\vec{r})]\cos(\omega_{12}\,t),\end{aligned} \tag{2.33}$$

which is an oscillation with a pulsatance $\omega_{12}\equiv\omega_1-\omega_2=(E_1-E_2)/\hbar$ correspondently with the Bohr's condition.

2.3 Schrödinger equation for a single particle

Though a flying particle can move along an arbitrary complicated trajectory, at a given time, it is in a well defined position. Waves have much more subtle patterns: during their propagation they can collide blocks and can be reflected from them. According to the superposition theorem, the sum of the incident and reflected waves become more and more sophisticated. How can a wave pattern like this be described?

Optics has already answered this question in the 19th century: we must not care for booking of the multiple reflections, rather it is advisable to loof for a *wave equation*: we should find a linear partial differential equation which can be satisfied by any of the plane waves separately and – because of linearity – their superpositions too. After finding the equation we have to solvi it with respect to the given boundary conditions. This guideline was determined and achieved by Schrödinger (1926) in quantum mechanics. Thus at that time de Broglie's playful fantasy about matter waves became a sterling and unbelievable effective part of theoretical physics.

2.3.1 The "deduction" of the Schrödinger equation

Let us find out the steps for Schrödinger equation. We have already known that a plane wave with a wave vector $\vec{k}$ and a pulsatance ω is the quantum mechanical analogy of moving particle with a momentum $\vec{p} = \hbar\vec{k}$ and a Eenergy $E = \hbar\omega$. Hence the amplitude of a plane wave (or in other word *wave function* as named by Schrödinger) can be written in the following way:

$$\Psi(\vec{r}, t) = e^{i(\vec{k}\cdot\vec{r} - \omega\, t)} = e^{\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - E\, t)}. \tag{2.34}$$

Moving in a force free space, the energy of the particle consists of kinetic energy only:

$$E = \frac{p^2}{2m}. \tag{2.35}$$

This condition restricts the possible shape of plane wave 2.34. The condition itself can be checked by *reading* the energy and the momentum appertaining to the plane wave 2.34 and putting them in equation 2.35. Our experiences related to the

superposition theorem suggest to deduce a linear equation, hence the operation of reading has to be linear too. Since in equation 2.34, energy and momentum are in the exponent, the adequate operation is the derivation of the wave function $\Psi(\vec{r},t)$. We will have a time-derivative and place vector-derivative (the latter means *gradient* denoted by the symbol $\vec{\nabla}$ (nabla)). Since in equation 2.35 p^2 is found, operation $\vec{\nabla}$ has to be done twice:

$$\begin{aligned} \frac{\partial}{\partial t}\Psi &= -\frac{i}{\hbar}E\Psi; \\ \vec{\nabla}\Psi &= \frac{i}{\hbar}\vec{p}\Psi; \\ (\vec{\nabla}\cdot\vec{\nabla})\Psi \equiv \triangle\Psi = \left(\frac{i}{\hbar}\right)^2(\vec{p}\cdot\vec{p})\Psi &= -\frac{1}{\hbar^2}\,p^2\Psi, \end{aligned} \tag{2.36}$$

where $\triangle = \vec{\nabla}\cdot\vec{\nabla}$ is the Laplace operator. After arranging these results:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\Psi &= E\Psi; \\ -i\hbar\vec{\nabla}\Psi &= \vec{p}\Psi; \\ -\frac{\hbar^2}{2m}\triangle\Psi &= \frac{p^2}{2m}\Psi. \end{aligned} \tag{2.37}$$

As we can see these equations are *eigenvalue-equations* of differenctial operators. Namely: energy E, momentum $\vec{p}$, and kinetic energy $p^2/2m$ can be read as the eigenvalues of – in sequence – the following operators: $i\hbar(\partial/\partial t)$, $-i\hbar(\vec{\nabla})$ (it is actually a vector-operator) and $-(\hbar^2/2m)\triangle$. As it has turned out, plane wave 2.34 is the common eigenfunction of these operators. Now we are ready to use the condition 2.35:

$$i\hbar\frac{\partial}{\partial t}\Psi = -\frac{\hbar^2}{2m}\triangle\Psi. \tag{2.38}$$

This was a significant step: we got the *Schrödinger equation* (which is a motion equation) of the free (force free) motion! From this moment we are not restricted by the plane wave picture or the eigenfunctions of energy and momentum, but we can look for their arbitrary superpositions in accordance with the initial boundary conditions.

There is one more important step we have to do: we need to handle the presence of a force field. In the simpliest (and commonest) case, there is a potential energy $V(\vec{r})$ depending on the position. In this situation, according to the logic of classical physics equation 2.35 has to be completed by inserting this potential energy:

$$E = \frac{p^2}{2m} + V(\vec{r}) = H(\vec{p}, \vec{r}), \tag{2.39}$$

where on the right side, H is the *Hamiltonian function* (known from classical mechanics), which – in case of a simple motion with a scalar potential– corresponds with the energy depending on position and momentum. As it can be seen, when energy E is a constant, momentum depends on position only. While $V(\vec{r})$ is a smooth function, consequances can be found out easily: though connection $\vec{p} = \hbar\vec{k}$ remains, the wave will not be a plane wave anymore, but a wave motion whose wavelength is changing from place to place, accordance with equation 2.21. Value of the wave vector can be read from place to place with the help of the operator $-in\vec{abla}$. From this, we get a wave eqquation:

$$i\hbar\frac{\partial}{\partial t}\Psi = -\frac{\hbar^2}{2m}\triangle\Psi + V(\vec{r})\,\Psi. \tag{2.40}$$

Though it is a crazy courage to try out how this equation works (as a rough mathematical model) if $V(\vec{r})$ is not a smooth function but it changes quickly in the range of the wavelength (just like in the hydrogen atom at the inner Bohr orbits), the result is shocking: the solutions of the equation yield the energy of the stationary electron orbits (and a lot of other properties discussed later) with a very high accuracy.

Thus equation 2.40 is not a rough model but a new essential equation of the physical description of nature. This is the *Schrödinger equation* of a particle (with a mass m) moving in a force field featured by a potential $V(\vec{r})$.

2.3.2 The quantum state and the Hamiltonian operator

The equation obtained at the end of the previous subsection points beyond the wave equation of a single particle. First af all, the motion of the "particle" is described as a wave motion which is much more various than we knew it in the picture of the classical mechanics. The *quantum state* of a particle can be described

by giving the complex function $\Psi(\vec{r}, t)$. The 2.40 Schrödinger equation describes the time evolution of the quantum state. This equation has a basic structure which can be written in the following way:

$$\frac{\partial}{\partial t}\Psi = -\frac{i}{\hbar}\hat{H}\Psi, \tag{2.41}$$

where $\hat{H}$ is the *Hamiltonian operator* whose shape (in the particular case) can be seen below:

$$\hat{H} = \frac{\hbar^2}{2m}\triangle + V(\vec{r}). \tag{2.42}$$

This operator can be built, if we have the H Hamiltonian function of the given system (for simple cases we need its energy expressed with momentums and coordinates) and we replace the momentum with the *momentum operator* met in equation 2.37:

$$\vec{p} \to \hat{\vec{p}} = -i\hbar\vec{\nabla}. \tag{2.43}$$

Henceforward operators which are the quantum mechanical analogy of some quantity in classical physics will be denoted by putting a ”hat” on its ”classical” notation. The role of the operators begins to stand out. Keeping at the particular case of the momentum, the ornated waves are mostly the superposition of plane waves propagating along a wide variety of directions. Hence in a motion like this, momentum does not have a definite value in general: plane waves – eigenfunctions of the momentum operator or in other word – refering to the physical content – *eigenstates* of the momentum – are just a kind of reference. However, in case of momentum measurement, this reference has a central role: at the beginning of the measurement, wave function are divided in space, according to modes limited by sideward. Though these modes are limited in space, they can be considered plane waves (the shape of them are almost the same). Commonly, this procedure means that the wave function is divided according to the approximate eigenstates of the quantity to be measured (in case of momentum this is done by an oprival lattice). Each detector selects a part wave (and only one detector belongs to each part wave) and the corresponsive eigenvalue is considered the result of the measurement.

In section 2.6 we will meet the mathematical analogy of the experimental division: any complicated wave function can be expressed according to the eigenfunctions

of the quantity to be measured (in case of momentum, this expression is just the Fourier serie) and the statistics of measurement of the given quantity can be calculated from the factors of the expression. The above mentioned construction of the Hamiltonian operator seems to be one-sided: it is not clear why we replaced the momentum with an operator and why we did not the same with the place coordinate. Actually we did it, just did not realise, because *the operator of the position is the multiplication with the space coordinate* which is valid for all function depending on place (for instance for $V(\vec{r})$): hence $V(\vec{r})\Psi$ is just the correct operator combination.

The explanation of this strange statement is that eigenfunctions of the position operator are functions localized on a definite place. Dirac delta function is the prototype of this kind of functions: it is an infinitely high and narrow barbs with a unit volume localized in some place $\vec{r}_0$: $\int \delta(\vec{r}-\vec{r}_0)d^3r = 1$; $\delta(\vec{r}-\vec{r}_0) = 0$ if $\vec{r} \neq \vec{r}_0$. If a position operator defined in the way above has an affect on a function like this, we get the following result:

$$\hat{\vec{r}}\,\delta(\vec{r}-\vec{r}_0) = \vec{r}\,\delta(\vec{r}-\vec{r}_0) = \vec{r}_0\,\delta(\vec{r}-\vec{r}_0), \tag{2.44}$$

so it works as it can be expected from an operator: it multiplies its eigenfunction by its eigenvalue.

2.3.3 The commutation relation of the position and the momentum

Knowing the operators of the position and the momentum we learn an essential and very important mathematical feature which can be used to solve a wide variety of problem in quantum mechanics without solving the Schrödinger equation. In addition to this it is used as a heuristic compass in a lot of cases when particular laws are not discovered yet. This feature is the following: if an operator acts on a function and then another one acts on the same function, the order of their action is not irrelevant in point of the result: operators are mostly *incommutable*.

Firstly, let us consider functions $f(x)$ depending the coordinate x only. If operators act on a function of them (one after another), it is obvious how to write these actions: it is the *product* of the operators. Operator which acts on the function first has to be written on the right and the second one has to be put on the left. According to the above mentioned, $x(\partial/\partial x)f(x) = xf'(x)$; $(\partial/\partial x)xf(x) = f(x) + xf'(x)$. The difference between them is $f(x)$. Symbolically it can be writ-

ten in the following way: $(\partial/\partial x)x - x(\partial/\partial x) = \hat{I}$ Now let us try to calculate the effect of the commutation of operators $\hat{p}_x = -i\hbar(\partial/\partial x)$ and $\hat{x} = x$. First of all it is useful to introduce the notation of the *commutator* of two operators (for example $\hat{A}$ and $\hat{B}$):

$$[\hat{A}, \hat{B}] =: \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{2.45}$$

Then according to the above mentioned:

$$[\hat{p}_x, \hat{x}] = -i\hbar. \tag{2.46}$$

Obviously, the same *commutation relation* applies to coordinate and momentum operators with directions either y or x. On the other hand $\hat{p}_x$ and y commutable, because in case of derivation according to x, y act like a constant. Summarized:

$$[\hat{p}_i, \hat{r}_j] = -i\hbar\delta_{ij}, \tag{2.47}$$

where i, $j = x$, y, z and $\delta_{ij} = 1$ if $i = j$, 0 otherwise (Kronecker-delta). These commutation relations were first introduced by Heisenberg, without reference to Schrödinger equation, using an abstract mathematical formalism.

The trivial generalisation of this result is the ”canonical quantization”: similar commutation relations are valid between canonically conjugated momentum – coordinate pairs. Nevertheless there are numerous exceptions of this rule, hence it is practical to calculate the commutators from directly the particular shape of the operators.

2.3.4 Stationary states and the time-independent Schrödinger equation

Now it is easy to define the quantum mechanical analogy of the stationary, non-radiating Bohr's orbits. These quantum states are such solutions of 2.41 Schrödinger equation where time dependence and place dependence split into a product shape and time dependence belongs to a defined frequency or in other word: a defined energy. The shape of these wave functions (which are the solutions of the equation) can be seen below:

$$\Psi(\vec{r},t) = e^{-i\omega t}\Phi(\vec{r}) = e-\frac{i}{\hbar}E\,t\Phi(\vec{r}). \tag{2.48}$$

Substituting this into the equation 2.41 and simplifying with $e^{-i\omega t}$, we get this equation:

$$\hat{H}\Phi = E\Phi, \tag{2.49}$$

which is extremly expressive: energy value $E = \hbar\omega$, which defines the ω oscillation frequency of the complex wave function is one of the eigenvalue of Hamiltonian operator $\hat{H}$. So *stationar states are energy eigenstates* and their sharply defined energies can be obtained by solving the eigenvalue problem belonging to the Hamiltonian operator. The name of equation 2.49 is *time independent Schrödinger equation* or *energy eigenvalue equation*. Probability density $|\Psi\vec{r},t)|^2$ belonging to the wave function whose shape is given in 2.48 does not change in time, thus the density of the charge distribution of the electron is a constant too. Most of the time a Hamiltonian operator is a differential operator. From this it follows that its eigenvalue problem (as a partial differential equation) remains undefined until boundary conditions referring to the eigenfunctions are not given. Indeed, though solutions to the equation existence for any E energy value, many of them does not satisfy the boundary conditions. Requirements of physics show up among the mathematical tools in the shape of the boundary conditions.

2.4 Properties of the solution of Schrödinger equation

The physical content of the Schrödinger equation is connected with direct mathematical properties. To know the physical background can be an aid in solving the mathematical problem. In this short section we will meet the physical properties of the finding probability of remaining (non emerging and non vanishing) particles: we will see that probability density is flowing from one place to another one as a fluid, but it never vanishes. In addition to this, we will review the typical cases of boundary conditions which define the energy eigenvalues.

2.4.1 Normalization

In subsection 2.2.8, we met Born's rule according to which the probability of finding a particle (whose state is $\Psi(\vec{r},t)$) in the volume d^3r near the position $\vec{r}$ at a time t equals $|\Psi(\vec{r},t)|^2 d^3r$. Finding at different places at the same time is not possible. This is called "particle properties" which is not a trivial feature but one of the essential experiences of quantum mechanics: in case of one piece of particle one detector may fire only. Hence detecting a particle *here* an detecting the same particle *there* are exclusive events. The ensemble probability of finding a particle here *or* there is the sum of the probabilities of the individual events. If a particle exists, it will be found somewhere by the detectors: its probability is the sum of the probabilities of all possible detection events:

$$\int |\Psi(\vec{r},t)|^2 d^3r = 1. \tag{2.50}$$

This expression is called the *normalization* of the wave function. Since Schrödinger equation is linear (both time-dependent and time-indepentent), it is no need to care about the normalization of the solution permanently, it is enough to multiply the obtained solution by a constant making 2.50 true. However, this can be done in only the case when integral 2.50 is finite. In this case wave function is said to be normalizable. Condition which makes wave function normalizable is often used as a boundary condition in solving the time-independent Schrödinger equation: discreet eigenvalues describing physical reality is selected from the continuum of energy values E by this condition many times. It is important to know that there are functions which can not be normalized among the simple wave functions which we have already met: one of these wave function is the plane wave with an infinit spacial extension. This is the eigenfunction of the momentum operator. The other is the eigenfunction of the operator of position vector, namely the Dirac delta function. Nevertheless we need to use these functions, thus their size has to be held. Fortunatelly there is an other way for it (more on this later).

2.4.2 The boundary conditions and the spectrum

Using the 2.49 Schrödinger equation with the shape of the Hamiltonian operator in expression 2.42, we get the following differential equation:

$$-\frac{\hbar^2}{2m}\triangle\Phi(\vec{r}) = (E - V(\vec{r}))\Phi(\vec{r}). \tag{2.51}$$

In this equation the place-dependent potential energy $V(\vec{r})$ is an input data. Not known things to be defined are energy eigenvalue E and its energy eigenfunction $\Phi(\vec{r})$. As we know this equation has solutions for any energy value E. We can select from these, if we base on solutions where $\Phi(\vec{r})$ satisfies boundary conditions required by physics. The whole bulk of solutions allowed by the boundary conditions is called the *spectrum* of the Hamiltonian operator.

The mentioned boundary conditions are the following

- the wave function has to be *continuous*.
- the wave function has to be *single-valued* function of the coordinates.
- in case of describing a *bound* state, the wave function has to be *normalizable*.

Spectrum defined by the boundary consitions can have both discreet and continuous regions. From the microworld, discreet specrum was the first message about natural laws differing from the classical rules. Typically, discreet energy eigenvalues can be observed in case of energy eigenstates bounded in a potential well (for instance electrons bounded in an atom). Several eigenfunction may belong to a discreet eigenvalue. This kind of eigenvalue is said to be *degenerate*. Free propagating waves have continuous energy eigenvalues.

2.4.3 Matter conservation and the complex wave function: the continuity equation

For the probability density $|\Psi(\vec{r},t)|^2 = \Psi(\vec{r},t)\Psi^*(\vec{r},t)$ an important conservation theorem follows from the wave equation 2.40 and its complex conjugate. The deduction:

$$\partial_t(\Psi\Psi^*) = \Psi^*\partial_t\Psi + \Psi\partial_t\Psi^* = \frac{i\hbar}{2m}(\Psi^*\triangle\Psi - \Psi\triangle\Psi^*). \tag{2.52}$$

Using the Green theorem (from vector analysis), we can rewrite this equation into the following form:

$$\partial_t(\Psi\Psi^*) = -\nabla\cdot\vec{j}, \tag{2.53}$$

where

$$\vec{j} = -\frac{i\hbar}{2m}(\Psi^* \nabla \Psi + \Psi \nabla \Psi^*) \tag{2.54}$$

is the density of the probability current. Equation 2.53 is a formal continutity equation, hence it describes the conservation of the total probability. This is the reason why in case of time-dependent Schrödinger equation, it is enough to normalize in the initial condition because it holds true afterwards.

2.5 A simple solution of the Schrödinger equation

In a standard course for physicist students, one begins to learn quantum mechanics in the following way: each student solves the time-independent Schrödinger equation for some simple cases. Probably there is no better method for a physicist student. Nevertheless, it is important to keep in mind we are not in a course like this and persons taking part in this course has never been trained in physics since they became university students. In addition to this, all the basic knowledge which help them understand quantum information has to be included in the opening of this subject. Hence – having few time – we have to skip or shorten many things that can be found in this introductory chapter. Such things are the simple solutions of the Schrödinger equation too. In a physiscist course this section would contain the description of the cases enumerated below

- bound states of an infinite potential well
- reflection symmetry
- finite potential
- particle locked up in a box
- tunnelling effect and resonance scattering
- stationary states of the harmonic oscillator

Because of the above-mentioned reasons, actually we should skip all this section. However, on the other hand, for the sake of presentation of the kind of thinking necessary in quantum mechanics, we need to show at least one of the cases enumerated above.

Let the harmonic oscillator be the presented system.

2.5.1 The stationary states of the harmonic oscillator

Harmonic oscillator is probably the most important *solvable* modell in physics. The word "solvable" means this modell can be totally followed by an analytic (not numeric) calculation. This modell is used to describe numerous systems found in nature: vibrations of molecules and cristalls, and – after some simple replacements of variables – vibrations of the electromagnetic field can be featured by this modell too. Besides the mentioned systems, there are a number of other systems that can be described if they are constructed by many harmonic oscillators. Furthermore there are systems that can be understood, if they are approximated by perturbated oscillators. In this subsection we will deal with this fundamental modell only. The treatment itself will be built in a simple way without using the general formalism of quantum mechanics[8].

Consider a particle with a mass m. A flexible force has an effect of his particle: $F = -Dx$. The solution of this equation is

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -Dx \tag{2.55}$$

$$x = x_0\cos(\omega t + \delta). \tag{2.56}$$

It is clear that the particle does vibrational motion along the direction of the force, where x_0 is the amplitude of the vibration, $\omega = 2\pi\nu = \sqrt{\frac{D}{m}}$ is the pulsatance and δ is a phase constant. A particle (or masspoint) with a harmonic vibrational motion is called linear oscillator. Now, we calculate the energy-eigenvalues of the oscillator. First of all, we consider its energy-operator:

$$\hat{E} = \frac{\hat{p}^2}{2m} + \frac{\omega^2 m}{2}\hat{x} \tag{2.57}$$

As it is known, in a case like this $\hat{p} = \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}x}$ and $\hat{x}$ corresponds to the multiplication by x. Using these knowledge, we write the eigenvalue-equation:

[8]Nevertheless we will learn about this formalism in the next section.

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}+\frac{\omega^2 m}{2}x^2\psi=E\psi \tag{2.58}$$

After reshaping it, we get the following expression:

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}+\frac{2m}{\hbar^2}(E-\frac{1}{2}m\omega^2x^2)\psi=0 \tag{2.59}$$

We are looking for the regular solutions of this equation and eigenvalues belonging to them. We can do it using *Sommerfeld*'s polinom method. Its essence is that first we find the asymptotic solution of the equation (for big x values) and then we multiply the obtained solution by a polynom with a finite power and look for the regular solution of the original equation in this shape. Let us introduce the following notation

$$k=\frac{2E}{\hbar\omega} \tag{2.60}$$

and a new variable (instead of x)

$$\xi=\sqrt{\frac{m\omega}{\hbar}}x\,. \tag{2.61}$$

Thus 2.59 gets a new form:

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}\xi^2}+(k-\xi^2)\psi=0. \tag{2.62}$$

The asymptotic shape of this equation can be seen below:

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}\xi^2}-\xi^2\psi=0. \tag{2.63}$$

Its solution is

$$\psi_a=e^{-\frac{\xi^2}{2}}. \tag{2.64}$$

The exact solution of 2.62 is sought in the following shape:

$$\psi = e^{-\frac{\xi^2}{2}} v(\xi). \tag{2.65}$$

If we put this expression into 2.62, we get a differential equation which defines $v(\xi)$:

$$\frac{\mathrm{d}^2 v}{\mathrm{d}\xi^2} - 2\xi \frac{\mathrm{d}v}{\mathrm{d}\xi} + (k-1)v = 0. \tag{2.66}$$

Function $v(\xi)$ is written in a polynomial shape:

$$v(\xi) = \sum_{r=0}^{n} c_r \xi^r. \tag{2.67}$$

Coefficients of the polynom has to be defined in such a way that the solution 2.65 can be a regular one. Let us create the first and second differential quotients of $v(\xi)$:

$$\frac{\mathrm{d}v}{\mathrm{d}\xi} = \sum_r r c_r \xi^{r-1}, \tag{2.68}$$

$$\frac{\mathrm{d}^2 v}{\mathrm{d}\xi^2} = \sum_r r(r-1) c_r \xi^{r-2}. \tag{2.69}$$

After putting them into 2.66, on the left side of the equation we get a series advancing according to the powers of ξ:

$$\sum_r \{(r+2)(r+1)c_{r+2} - (2r+1-k)c_r\}\xi^r = 0. \tag{2.70}$$

This equation holds true for all values of ξ, if the coefficients of all powers of ξ equal 0. After vanishing of the coefficient of ξ^r, we get a connection between the coefficients of c_{r+2} and c_r:

$$c_{r+2} = \frac{2r+1-k}{(r+2)(r+1)} c_r. \tag{2.71}$$

It can be seen that function 2.65 will be regular in the case only when polynom $v(\xi)$ has a finite power, in other words from a particular $r = n$ every coeffitients are identically zero. If the greatest exponent is denoted by n, then – according to 2.71 – $c_{n+2} = 0$. It comes true, if

$$2n + 1 = k. \tag{2.72}$$

From the connection 2.60 between eigenvalue E and parameter k, we get:

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) = h\nu\left(n + \frac{1}{2}\right). \tag{2.73}$$

As we can see, energy-eigenvalues of the oscillator can not have arbitrary values, because they may be the integral multiples of the quantum $h\nu$ only. The eigenvalue which belongs to $n = 0$ is

$$E_0 = \frac{h\nu}{2}. \tag{2.74}$$

This is called the *zero point energy* of the oscillator. This very interesting fact means that a quantum mechanical oscillator has a nonzero sum of energy, even in its totaly relaxed state(!). In addition to this: a quantum mechanical oscillator may not have an energy which is less than the calculated zero point energy. In classical physics, it is not possible. Difference of two neighbouring levels of energy:

$$E_n - E_{n-1} = h\nu. \tag{2.75}$$

As we saw, using operators quantum mechanics gives a natural explanation of the *Planck*-hypothesis.

The eigenfunction which belongs to the eigenvalue E_n is

$$\psi_n = e^{-\frac{\xi^2}{2}} v_n(\xi). \tag{2.76}$$

The polynom $v_n(\xi)$ is called the n-th degree *Hermite*-polynom and it is denoted by $H_n(\xi)$. Based on 2.66, the differential equation defining $H_n(\xi)$ can be written as

$$\frac{\mathrm{d}^2 H_n}{\mathrm{d}\xi^2} - 2\xi\frac{\mathrm{d}H_n}{\mathrm{d}\xi} + 2nH_n = 0. \tag{2.77}$$

So the *n*-th eigenfunction is

$$\psi_n = e^{-\frac{\xi^2}{2}} H_n(\xi). \tag{2.78}$$

These results can be applied for explaining the vibrational spectrum of mulecules that consist of two atoms, because in such molecules atoms are bounded by forces that are nearly flexible:

$$F = -D(r - r_0), \tag{2.79}$$

where r is the distance between the nuclei of the atoms and r_0 is their equilibratory distance. By reason of the force 2.79, molecule does flexible vibrations whose possible energy-values are given by eigenvalues 2.73. When a molecule goes from an excited state to a state that belongs to a lower level of energy, it emits a radiation whose frequency can be calculated by using *Bohr*'s frequency-condition:

$$\nu = \frac{E_n - E_{n-1}}{h} = \frac{1}{2\pi}\sqrt{\frac{D}{m}} \tag{2.80}$$

After measuring the frequency of the emitted light, coefficient D of the flexible force that holds atoms together can be calculated by using 2.80. For example, for molecule of hydrochloric acid, $D = 4.806 \times 10^2 \mathrm{N}\cdot\mathrm{m}^{-1}$ is given by measurements. The existence of zero-point-energy has been proved by observations in molecular spectroscopy.

Before finishing this section, it is important to introduce a quantity which belongs to this topic but will be useful later. Though the calculation of the normalization integral of Hermite polynom is not a trivial task, eigenfunction $n = 0$ can be obtained based on knowledges in probability theory: the zeroth-order is a constant, hence the relating probability distribution function is a Gauss distribution with some scatter: $|\psi_0(x)|^2 = e^{(-x^2/2\sigma^2)/(\sqrt{2\pi}\sigma)}$, that is:

$$\psi_0(x) = \frac{1}{\sqrt{\sqrt{2\pi}\sigma}} e^{-\frac{x^2}{4\sigma^2}}. \tag{2.81}$$

Compared this to the function shape $e^{-\frac{\xi^2}{2}}$ and the definition of ξ, scatter of the ground state wave packet can be read:

$$\sigma = \sqrt{\frac{\hbar}{2m\omega}}. \tag{2.82}$$

It is worth memorizing this result. Let us notice that scatter becomes 0 – in accordance with classical mechanics – if $\hbar$ vanishes or the mass of the oscillator is very big or in case of a very strong spiral.

There is another thing that has to be mentioned, namely the energy eigenvalue belonging to $n = 0$. As we saw in this case energy is not 0 but $E_0 = h\nu/2 = \hbar\omega/2$ (the zero point energy). As it was dicussed in subsection 2.2.7, the available minimum energy is ensured by not the localization on the bottom of the potential well, but a wave packet with a Gaussian curve whose width is $\sqrt{\hbar/2m\omega}$. This wave packet has both a little potential and kinetic energy, but their sum is the smallest.

2.6 The general formalism

The region of physics which had been known before quantum mechanics was born is named *classical physics* and considered as a simplified picture of our world. However this simplified picture gives an amazingly precise description of the macroscopic world and it has been never caught at any laxaty, nowhere. In classical physics the state of a pointlike body (or particle) at a time t is given by two vectors: one of these denotes the position of the body, the other one gives its speed. If we know the forces acting on the body and the 6 components of the mentioned vectors, we can calculate (predict) how the particle will move in the future. This rigorous causality one of the best proved facts of classical physiscs: this is the reason why planning of our houses and machines are entrusted to engineers, insdead of magician.

In microscopic world, we need a more sophisticated and detailed description: a structureless, pointlike particle moves spaciously as a wave and wave function $\Psi(\vec{r}, t)$ which describes the amplitude of the wave motion changes in time. And this function describes the state of the particle at a time t. It has all information about position, momentum and anything which we want to know. The time evolution of this function described by Schrödinger equation which is as strict law of nature as Newton equations (which describes the motion of a classical body).

We have already met the randomness which appeared at the measurement of quantum systems. Born's rule dealt with the measurement of the position only. This will be generalized to measurements of other quantities now. The obtained statistic laws describe a wide variety of experiments with different systems.

A long time ago, a statistical approach of multipartite systems was possible only. However, in the last decades, the formidable development of experimental technology on the level of individual physical systems has made the direct observation of this kind of objects viable.

During the discovery of the consequences of the Schrödinger equation, a beautiful mathematical structure appears: this is the *Hilbert space* whose vectors are quantum states moved by operators belonging to physical quantities. Numerous results will be obtained from the compact expressiveness of this picture. This section is aimed to collect them.

2.6.1 Wave function and quantum state; observable physical quantities and operators

Particles with a state described by the wave function generally do not have definite position, momentum, angular momentum, energy, except for the special cases when wave function is the eigenfunction of the operator of the particular physical quantity. We have already met some of these cases. We have learned that $\delta(\vec{r}-\vec{r}_0)$ is the wave function of a particle localized on the place $\vec{r}_0$. On the other hand $\delta(\vec{r}-\vec{r}_0)$ is the eigenfunction of the *position operator*[9] for its eigenvalue $\vec{r}_0$. Similarly, $e^{(i/\hbar)\vec{p}\cdot\vec{r}}$ is the wave function of a particle with a momentum $\vec{p}$. On the other part this is the eigenfunction of the *momentum operator*[10] for its eigenvalue $\vec{p}$.

The global rule is the following: Each observable and measurable physical quantity has an operator whose eigenfunctions are the wave functions in which the given quantity has a definite value. This definite value is just the eigenvalue of the operator and belongs to the given eigenfunction.

Most of the time, a wave function is not an eigenfunction of an operator of some known physical quantity. Nevertheless, it can be constructed as a linear combination of the eigenfunctions of an arbitrary physical quantity. Mathematically, Fourier line extraction of a function means exactly this. Physically, as we will see

[9] As we know, its operation is the multiplication with $\vec{r}$.

[10] Its operation is $(\hbar/i)\vec{\nabla}$.

in the next point, a construction like this is closely linked to the measurement of the given quantity. Eigenfunctions of operators of the essential physical quantities form a complete system. This means that any kind of sophisticated wave function can be constructed with their linear combination.

If each eigenfunction of two operators are mutual, then the operators are commutable. Indeed, if $\hat{A}\Phi_n = a_n\Phi_n$ and $\hat{B}\Phi_n = a_n\Phi_n$, then $(\hat{A}\hat{B} - \hat{B}\hat{A})\Phi_n \equiv [\hat{A}, \hat{B}]\,\Phi_n = 0$. Hence $[\hat{A}, \hat{B}]\,f(\vec{r}) = 0$ for each function which can be written as the linear combination of the eigenfunctions. Because of the above-mentioned completness, this is true for every function, hence $\hat{A}$ and $\hat{B}$ are really commutable under every circumstances: $[\hat{A}, \hat{B}] = 0$.

Also the reverse statement is true: if $\hat{A}$ and $\hat{B}$ are commutable and each of them a complete eigenfuction system separately, then their eigenfunctions are mutual or (in case of degenerate) mutual linear combinations can be selected from them.

A typical example for non commuting operators is the case of place vector and the momentum. Eigenfunctions of the place vector are the delta function localized on a point and eigenfunctions of the momentum are plane waves covered the whole space. These are as different as two things can differ from each other.

Operators assigned for physical quantities are the basic tools of quantum mechanics. Their role is not restricted to the selection of eigenstates and eigenvalues. Though due to the superposition theorem, quantum states are not the eigenstates of a physical quantity in general, operators are sterling mathematical expressions of the given quantity in both the dynamics of the quantum state and the measurement of the quantity itsels. The latter is an especially hot topic in quantum mechanics and we will begin to learn it in the following subsection.

2.6.2 The statistics of the measurement

For a special case, we have already known how the wave function defines the statistics of the measurement: this is the measurement of the position. The place vector of the particle is measured by detectors with a small enough size. It is known that one piece of particle can make only one detector fire. The detector (which fires) is selected by a random event. The position (or place vector) of this detector is the result of the position measurement. According to Born's rule $|\Psi(\vec{r},\, t)|^2\, d^3r$ is the probability of finding the particle around a place denoted by $\vec{r}$, in a volume d^3r by a detector. Since detector-clicks in different places are mutually exclusive events, the average value of the place vector can be obtained

via a direct integral:

$$\langle \vec{r} \rangle = \int \vec{r} |\Psi(\vec{r},\, t)|^2 \, d^3r = \int \Psi^* \hat{\vec{r}} \Psi d^3r, \tag{2.83}$$

where – in the last step – we use that position operator is the multiplication with the place vector. How does work the statistics of randomness in case we measure another quantity, for example the momentum? Since the momentum equals $\hbar$ times the wave vector, first of all it is evident to dissolve the matter wave with an optical lattice according to the wave vectors. Components (the eigenfunctions of the momenzum operator with different wave vectors) are swung towards different directions and in case the distance from the lattice is far enough, they are separated in space. At a distance like this, path of each component is closed by a detector. The result of the measurement is the momentum which belongs to the wave vector whose path is closed by the detector that fires. The question is how to define the probability of measuring a given momentum value.

The answer will be sought in a more general way in case of the measurement of a physical quantity A. Consider $\hat{A}$ which is the quantum mechanical operator belonging to the physical quantity A. Let us find its eigenvalues and normalized eigenfunctions:

$$\hat{A}\varphi_n(\vec{r}) = a_n \varphi_n(\vec{r}). \tag{2.84}$$

Beacuse of the operating way of the "A separator" device, eigenfunctions of $\hat{A}$ are swung to different places:

$$\varphi_n(\vec{r}) \Rightarrow \Phi_n(\vec{r}). \tag{2.85}$$

Functions Φ_n-s are normalized wave modes even now and do not overlap, because they are taken far enough from each other in order that they can not overlap. As János Neumann – who was a Hungarian-American scientist (1903-1957) and this strategy is named after him "von Neumann measurement" – suggested: to reach this, we have to use a force which depends on the quantity A. For example, in case of measurement of the momentum an optical lattice is applied, and in case of magnetic momentum an inhomogeneous magnetic field is used (Stern-Gerlach experiment).

Each space-separated Φ_n wave mode is covered by one of the detectors. In this situation, if our system was in a $\varphi_n(\vec{r})$ pure state before separating, the nth detec-

tor will certainly fire (with a probability value 1) and the other ones will certainly be calm.

However the quantum state of a particle is not equivalent to one of the eigenfunctoins in general. Most of the time, the wave packet (on which we want to measure the value of A) is built from several eigenfunctions. Since the separation is achived according to the linear Schrödinger equation, weights in the superposition remain unaltered:

$$\Psi(\vec{r}, t = 0) = \sum_n c_n \varphi_n(\vec{r}) \Rightarrow \sum_n c_n \Phi_n(\vec{r}). \tag{2.86}$$

In this situation, separated components do not interfere with each other, hence the probability value 1 is multiplied with $|c_n|^2$, according to Born's rule. So the probability of obtaining eigenvalue a_n is the following:

$$p_n = |c_n|^2. \tag{2.87}$$

This is one of the fundamental connections of quantum mechanics which is often said to be a postulate. Here we realized that the result follows from the strategy of spacial separation. However, fortunatelly, probabilities do not depend on the technical detailes of the particular measumerent, because they depend on factors c_n-s – which feature the state of the initial wave function – only. Normalization of the wave function survives the separation: we will certainly find the particle somewhere, thus

$$\sum_n p_n = \sum_n |c_n|^2 = 1, \tag{2.88}$$

which follows from the orthogonality of the eigenfunctions (see the next point).

The result in 2.87 defines the average value of measurement of the physical quantity A (in other words: the operator $\hat{A}$) in the state $\Psi = \sum_n c_n \phi_n$:

$$\langle \hat{A} \rangle_\Psi = \sum_n p_n a_n = \sum_n |c_n|^2 a_n. \tag{2.89}$$

2.6.3 The Hilbert space of the states; completeness; the quantum mechanical average

To express a function according some function system compares to the expression of a vector according to a system of unnit vectors (a "base"). In both case we need a basic operation, namely the *scalar product* which assignes a number to two vectors (or functions) to define the factors of the expression. In the simpliest case base vectors $\vec{i}$, $\vec{j}$ and $\vec{k}$ are *orthogonal*, that is the scalar product of any two base vaectors is 0 and they are *normalized* which means that the scalar product of any vector with itself is 1. A base with these two properties is said to be *orthonormal*. In this case, in an expression like this $\vec{a} = a_i\vec{i}+a_j\vec{j}+a_k\vec{k}$, factor a_j can be obtained in the following way: $\vec{j}\cdot\vec{a} = a_i\vec{j}\cdot\vec{i}+a_j\vec{j}\cdot\vec{j}+a_k\vec{j}\cdot\vec{k} = a_i\times 0+a_j\times 1+a_k\times 0 = a_j$. How can we define the scalar product in case of two functions? Normalization of finding probability can help us: connection $\int |\Psi|^2 d^3r = \int \Psi^*\Psi d^3r = 1$ can match to the scalar product of a unit vector with itself. Hence the definition of the scalar product between two complex functions $f(\vec{r})$ and $g(\vec{r})$ is defined in the following way:

$$\langle f|g\rangle = \langle g|f\rangle^* = \int f^*(\vec{r})g(\vec{r})d^3r. \tag{2.90}$$

With this definition, we introduced the *Hilbert space* of the quantum states featured with the wave function. The expressive and efficient structure of the Hilbert space serves to describe geometrically the set of elements (here: quantum states of a physical system) amoung which two operations can be defined: 1. linear combination, 2. scalar product. The first one is a linear operation ensured by the superposition theorem in quantum physics, and the second one is ensured by the quadratic Born rule and the numerous appearance of the resolution according to the eigenfunctions.

The right side of equation 2.90 is the definition of the scalar product, but now we focus on its left side where we introduced the famous "bra-ket" notation of Dirac. In this the great artifice is that the scalar product can be decomposed: if function $g(\vec{r})$ is considered as one of the vectors of the Hilbert space, its notation is $|g\rangle$. If the first coefficient of a scalar product is a function $f^*(\vec{r})$, it can be considered as one of the vectors of the "adjoint space" and its vector notation is $\langle f|$. Hence their scalar product is denoted by the following symbol: $\langle f|g\rangle$. Vectors denoted by the symbol $\langle\ |$ are named *bra vectors* and they are the adjoint vectors of those which are denoted by $|\ \rangle$, *ket vectors*.

Every possible state of a physical system is described by a ket vector. Bra vectors appear in scalar products only. Each ket vector $|a\rangle$ has a bra vector $\langle a|$ which is its adjoint. The adjoint of a ket $\alpha|a\rangle + \beta|b\rangle$ is the bra $\alpha^*\langle a| + \beta^*\langle b|$. This notation emphasize that the physical content of a quantum state does not depend on the used "coordinate system" in which the state is described. For example, a quantum state can be represented as the wave function of the place vector or with the Fourier coefficients of an expression according to the momentum-eigenfunctions or by chance as an expression of other base functions. Accordingly, in general, just an index is written within the symbols "bra" and "ket" as an identifier of the particular quantum state.

Let $|n\rangle$ $(n = 1, 2, ...)$ be the vector notation of an orthonormal base which consists of wave functions $u_n(\vec{r})$:

$$\langle m|n\rangle = \delta_{mn}, \tag{2.91}$$

and let us express the quantum state $|\Psi(t)\rangle$ – which matches the wave function $\Psi(\vec{r}, t)$ – according to the base vectors:

$$|\Psi(t)\rangle = \sum_n c_n(t)|n\rangle. \tag{2.92}$$

If we multiply this with the bra $\langle m|$ from the left, and we use the orthonormality of the base vectors, we get the coefficients of the expression:

$$|c_m(t) = \langle m|\Psi(t)\rangle \quad \left(\equiv \int u_m^*(\vec{r})\Psi(\vec{r}, t)d^3r \right). \tag{2.93}$$

Replacing this back to the expression 2.92, we get this:

$$|\Psi(t)\rangle = \sum_n |n\rangle\langle n|\Psi(t)\rangle. \tag{2.94}$$

Let us notice what it means:

$$\sum_n |n\rangle\langle n| = \hat{I}. \tag{2.95}$$

There is a nontrivial thing here, namely: there are *enough* base vectors $|n\rangle$ to express an arbitrary state Ψ. If this condition is satisfied, the orthonormal base is *complete*, and the name of the equation 2.95 is *completness relation*.

Transiently, let us return to the wave functions. What is the completness relation in their case? Using the definition of the scalar product, let us express the wave function according to a complete orthonormal system of functions:

$$\Psi(\vec{r},t) = \sum_n u_n(\vec{r}) \int u_n^*(\vec{r}\,')\Psi(\vec{r}\,',t)d^3r'. \tag{2.96}$$

From this, completness relation of wave function can be read:

$$\sum_n u_n(\vec{r})u_n^*(\vec{r}\,') = \delta(\vec{r}-\vec{r}\,'), \tag{2.97}$$

which is in accordance with equation 2.95.

Members $|n\rangle\langle n|$-s in the sum of equation 2.95 have an expressive and important meaning: these are *projector operators* (or in short: projectors) projecting onto base vectors $|n\rangle$. In general a projector $|\Psi\rangle\langle\Psi|$ acts on a ket vector which stands on its right side in the following way: firstly it multiplies the ket with $\langle\Psi|$ (this is their scalar product), this means it reads the length of the projection on the unit vector $|\Psi\rangle$. After this, it multiplies the same unit vector with the obtained number, that is it creates the projection.

In the important case when the elements of a complete orthonormal system $|n\rangle$ are the eigenvectors of an operator $\hat{A}$ which belongs to a physical quantity A and eigenvalue a_n belongs to the eigenvector $|n\rangle$, operator $\hat{A}$ acts on the state in the following way:

$$\hat{A}|\Psi(t)\rangle = \hat{A}\sum_n |n\rangle\langle n|\Psi(t)\rangle = \sum_n a_n|n\rangle\langle n|\Psi(t)\rangle = \equiv \sum_n |n\rangle a_n\langle n|\Psi(t)\rangle, \tag{2.98}$$

hence operator $\hat{A}$ can be symbolically written in the following shape:

$$\hat{A} = \sum_n |n\rangle a_n\langle n|, \tag{2.99}$$

which is named the *unit resolution* of the operator. This result can be used to written the ultimate shape of the expression of the average value of an operator:

$$\langle \hat{A} \rangle_\Psi = \sum_n |c_n|^2 a_n = \sum_n \langle \Psi | n \rangle \langle n | \Psi \rangle a_n = \langle \Psi | \sum_n |n\rangle a_n \langle n | \Psi \rangle, \tag{2.100}$$

that is

$$\langle \hat{A} \rangle_\Psi = \langle \Psi | \hat{A} | \Psi \rangle. \tag{2.101}$$

This is the essential average form on which most of the aplication of quantum mechanics is based. This can be generalized further when the average is further averaged to the statistic ensemble of quantum states. We will discussed it soon.

2.6.4 Self-adjoint operators

In quantum mechanics, most of the measurable physical quantities are inherited from the classical physics. Their measured values are obviously real numbers. The properties of *self-adjointness* ensures the eigenvalues of particular operators to be real. It has some important consequences which will be understable, if we learn the idea of a *adjoint operator*.

First of all, let us introduce the following notation: $\hat{A}|u\rangle =: |\hat{A}\, u\rangle$. In this case the adjoint of this vector is $\langle \hat{A}\, u|$. Using these notations, the adjoint operator $\hat{A}^\dagger$ of operator $\hat{A}$ is an operator which satisfies the following equation (with arbitrary vectors $\langle u|$ and $|v\rangle$):

$$\langle \hat{A}^\dagger u | v \rangle = \langle u | \hat{A}\, v \rangle. \tag{2.102}$$

An operator $\hat{A}$ is *self-adjoint* or in other word *hermitian*, if

$$\hat{A}^\dagger = \hat{A}. \tag{2.103}$$

Eigenvalues of a self-adjoint operator like $\hat{A}$ are real. It can be easily admitted by considering the properties written in expression 2.90:

$$a_n = \langle n|\hat{A}\, n\rangle = \langle \hat{A}\, n|n\rangle^* = \langle n|\hat{A}\, n\rangle^* = a_n^*, \tag{2.104}$$

which is really a real number. Self-adjoint operators have another important feature, namely: their eigenfunction belomging to different eigenvalues are orthogonal. Indeed, if $\hat{A}|n\rangle = a_n|n\rangle$ and $\hat{A}|m\rangle = a_m|m\rangle$, then

$$\begin{aligned} \langle m|\hat{A}\, n\rangle &= a_n\langle m|n\rangle \\ = \langle \hat{A}\, m|n\rangle &= a_m\langle m|n\rangle. \end{aligned} \tag{2.105}$$

From this, it follows that

$$(a_m - a_n)\langle m|n\rangle = 0, \tag{2.106}$$

that is there are two cases: one of them is the situation when $a_m = a_n$, the other is the orthogonality: $\langle m|n\rangle = 0$. Let is find if operators we (met so far) are really self-adjoint ones. For the multiplication with position coordinates, self-adjointness is obviously true. What is with the momentum operator? For the sake of simplicity, let us stay in one dimesion only: is it true that

$$\int u^*(x)\left(\frac{\hbar}{i}\frac{d}{dx}v(x)\right)dx = \int \left(\frac{\hbar}{i}\frac{d}{dx}u(x)\right)^* v(x)dx? \tag{2.107}$$

The answer is yes, because after a partial integral sides ov this expression transform into each other, provided functions $u(x)$ and $v(x)$ disappear on remote bounds or they satisfy a cyclic boundary condition. Both examples occur in quantum mechanics.

2.6.5 Operators and matrices; untitary transformations

We know everything about an opperator $\hat{A}$, if we know how it acts on an arbitrary state $|\Psi\rangle$. Our strategy is the following: with the help of formulae 2.92 and 2.93, we will express both the initial state $|\Psi\rangle$ and the vector $\hat{A}|\Psi\rangle$ (which came into being after the effect of operator $\hat{A}$) on an orthonormal base $|n\rangle$ ($n = 1, 2, 3, ...$). This operation can be written in such a way as to insert the unit operator (see the formula 2.95) before and after operator $\hat{A}$:

$$\hat{A}|\Psi\rangle = \sum_m b_m|m\rangle = \sum_m |m\rangle b_m$$
$$= \sum_m |m\rangle\langle m|\hat{A}|\Psi\rangle = \sum_m\sum_n |m\rangle\langle m|\hat{A}|n\rangle\langle n|\Psi\rangle = \sum_m |m\rangle\sum_n A_{mn}c_n, \tag{2.108}$$

where we introduced the matrix element $\langle m|\hat{A}|n\rangle \equiv A_{mn}$ of the operator $\hat{A}$. The readable result is

$$b_m = \sum_n A_{mn}c_n, \tag{2.109}$$

that is on the base $|n\rangle$ of the Hilbert space, the coefficient vector b_m which illustrates $\hat{A}|\Psi\rangle$ can be obtained from the coefficient vector c_n which illustrates $|\Psi\rangle$ by a multiplication with the matrix A_{mn}. Let us note that the matrix elements of a self-adjoint operator have the following simmetry:

$$A_{nm} = A^*_{mn}. \tag{2.110}$$

As a task, let us rewrite the Schrödinger equation into a matrix shape:

$$\partial_t|\Psi\rangle = -\frac{i}{\hbar}\hat{H}|\Psi(t)\rangle = -\frac{i}{\hbar}\hat{H}\sum_n c_n(t)|n\rangle. \tag{2.111}$$

If we multiply this with the bra $\langle m|$ from the left, we will get the sought matrix Schrödinger equation:

$$\dot{c}_m = -\frac{i}{\hbar}\sum_n H_{mn}c_n. \tag{2.112}$$

With this, let us write the adjoint of the time-dependent Schrödinger equation too:

$$\dot{c}^*_m = \frac{i}{\hbar}\sum_n H^*_{mn}c^*_n = \frac{i}{\hbar}\sum_n c^*_n H_{nm}, \tag{2.113}$$

which is perhaps a not too elegant way to prove the following often used formula:

$$\partial_t \langle \Psi | = \frac{i}{\hbar} \langle \Psi | \hat{H}. \tag{2.114}$$

The matrix representation of the operators yields a very flexible mathematical possibility to solve quantum mechanical problems with the tools of linear algebra using always the most appropriate base for the given problem.

Introducing the Hilbert space, we emphasized that meaning of the state vectors and operators does not depend on the way of their representations (the chosen base). Hence one of the basic operation in quantum mechanics is the *change of basis*. Let us perform this operation with the vectors and matrices above. Let the initial base be $|n\rangle$ and the final – orthonormal and whole – base be $|\alpha\rangle$. In other words let us find the factors d_α of the following resolution:

$$|\Psi\rangle = \sum_m c_m |m\rangle = \sum_\alpha d_\alpha |\alpha\rangle. \tag{2.115}$$

The calculation can be done by inserting the unit resolution:

$$d_\alpha = \langle \alpha | \Psi \rangle = \sum_m \langle \alpha | m \rangle \langle m | \Psi \rangle = \sum_m U_{\alpha m} c_m, \tag{2.116}$$

where we introduced the transformation matrix:

$$U_{\alpha m} := \langle \alpha | m \rangle. \tag{2.117}$$

What happens with the matrices of the operators associated measurable physical quantities? Naturally – as an effect of base change – they transform too:

$$\begin{aligned} \langle \alpha | \hat{A} | \beta \rangle &= \sum_{m,n} \langle \alpha | m \rangle \langle m | \hat{A} | n \rangle \langle n | \beta \rangle \\ &= \sum_{m,n} U_{\alpha m} A_{mn} (U^{-1})_{n\beta} = \left(\underset{= = =}{U A U^{-1}} \right)_{\alpha\beta}, \end{aligned} \tag{2.118}$$

where we used that $\langle n | \beta \rangle$ is the matrix element of U^{-1} inverse transformation which traces back to the initial base $|n\rangle$ from the "new" one $|\beta\rangle$.

The matrix U has a special mathematical feature: it is a *unitary* matrix, that is its inverse equals its adjoint: $U^\dagger U = \hat{I}$. This can be seen from the following expression:

$$U^{-1}_{m\alpha} = \langle m|\alpha\rangle = \langle \alpha|m\rangle^* = (U_{\alpha m})^*. \tag{2.119}$$

This is not a mathematical accident. Unitarity is the necessary and satisfactory condition for reserving the norm of quantum state under transformations (described by a linear operator) or base changes. Indeed – using base-independent operators – , if the unitarity condition is satisfied, that is

$$\hat{U}^\dagger \hat{U} = \hat{I}, \tag{2.120}$$

then the norm of the transformed state vector

$$|\Psi'\rangle = \hat{U}|\Psi\rangle \tag{2.121}$$

is

$$\langle \hat{U}\,\Psi|\hat{U}\,\Psi\rangle = \langle \Psi|\hat{U}^\dagger \hat{U}|\Psi\rangle = \langle \Psi|\Psi\rangle = 1. \tag{2.122}$$

Let us mark the following rule: if no measurement is made on a quantum system, every change in its quantum state can happen in a unitary way only. During a change like this, operators transform according to the following rule:

$$\hat{A}' = \hat{U}\,\hat{A}\,\hat{U}^\dagger, \tag{2.123}$$

as it can be read from the identity: $\hat{U}\,\hat{A}|\Psi\rangle = (\hat{U}\hat{A}\hat{U}^\dagger)|\Psi'\rangle$.

It is important to mention the time evolution of the state vector: since in this evolution the total probability remains unchanged, the time evolution is described by a unitary operator.

2.6.6 Continuous matrices; coordinate and momentum representation

As it is known $\delta(\vec{r}-\vec{r}_0)$ eigenfunctions of the coordinate operator (whose spectrum is continuous) can not be normalized but they are complete, because summarizing (integrating) them for the spectrum of eigenvalues $\vec{r}_0$, they satisfy the 2.97 completness relation:

$$\int \delta(\vec{r}-\vec{r}_0)\,\delta(\vec{r}\,'-\vec{r}_0)\,d^3r_0 = \delta(\vec{r}-\vec{r}\,'). \tag{2.124}$$

Since Dirac's delta function is real, also the orthonormal relation is replaced with this equation.

In Dirac notation, $|\vec{r}_0\rangle$ corresponds to the $\delta(\vec{r}-\vec{r}_0)$ wave function and the expression of the state vector – according to these coordinate-eigenvector – can be written into the following form:

$$|\Psi(t)\rangle = \int C(\vec{r}_0,t)\,|\vec{r}_0\rangle\,d^3r_0, \tag{2.125}$$

where the expression factor is

$$C(\vec{r}_0,t) = \langle\vec{r}_0|\Psi(t)\rangle = \int \delta(\vec{r}-\vec{r}_0)\,\Psi(\vec{r},t)\,d^3r = \Psi(\vec{r}_0,t), \tag{2.126}$$

that is, in "Hilbert space language", the "wave function" is actually the continuous vector of the expression factors of state vector on the base of eigenvectors of the coordinate.

The other continuous spectrum operator is the momentum, whose values – in case of boundless motion – can change continuously. In this situation, the expression according to the momentum-eigenfunctions will be a Fourier integral instead of a Fourier serie. Its fundamental form is

$$\int_{-\infty}^{\infty} e^{ikx}dx = 2\pi\,\delta(k), \tag{2.127}$$

where if the number of dimensions is greater than 1, in the exponent a scalar product has to be written and on the right side a $(2\pi)^d$ multiplier is needed. From

this we can obtain the whole system of the momentum-eigenfunctions belonging to the momentum-eigenstates $|\vec{p}\rangle$.

We look for them in the following shape: $C\, e^{\frac{i}{\hbar}\vec{p}\cdot\vec{r}}$ and normalization factor C can be defined from the condition that the following normalization connection has to be true for continuous spectrum:

$$\langle\vec{p}\,'|\vec{p}\rangle = \int \langle\vec{p}\,'|\vec{r}\rangle\langle\vec{r}\,|\vec{p}\rangle d^3r = \int C^*Ce^{\frac{i}{\hbar}(\vec{p}-\vec{p}\,')\cdot\vec{r}}d^3r = \delta(\vec{p}-\vec{p}\,'). \tag{2.128}$$

This connection plays the role of the completness relation too. Using the formula 2.127 and the feature according to which $\delta(x)/|a|$ and considering C a real number, we get the following result:

$$\langle\vec{r}\,|\vec{p}\rangle = (2\pi\hbar)^{-3/2}e^{\frac{i}{\hbar}\vec{p}\cdot\vec{r}}, \tag{2.129}$$

where the exponent of the normalization factor – in case of d dimensions – equals $d/2$. The quantum state – similarly as we can see in equations 2.125 and 2.126 – can be expressed via the eigenvectors of the momentum:

$$|\Psi(t)\rangle = \int \Phi(\vec{p},t)\,|\vec{p}\rangle\, d^3p, \tag{2.130}$$

where the factor

$$\begin{aligned}\Phi(\vec{p},t) &= \langle\vec{p}\,|\Psi(t)\rangle \\ = \int \langle\vec{p}\,|\vec{r}\rangle\langle\vec{r}\,|\Psi(t)\rangle\, d^3r &= \int \frac{e^{-(i/\hbar)\vec{p}\cdot\vec{r}}}{(2\pi\hbar)^{3/2}}\Psi(\vec{r},t)\, d^3r\end{aligned} \tag{2.131}$$

is the *wave function in momentum-representation*. This equation can be considered as the formula of the conversion from the coordinate-representation ($\Psi(\vec{r},t)$) to the momentum-representation ($\Phi(\vec{p},t)$). Using equations 2.129 and 2.130 it is easy to write the formula of reversion:

$$\begin{aligned}\Psi(\vec{r},t) &= \langle\vec{r}\,|\Psi(t)\rangle \\ = \int \langle\vec{r}\,|\vec{p}\rangle\langle\vec{p}\,|\Psi(t)\rangle\, d^3r &= \int \frac{e^{(i/\hbar)\vec{p}\cdot\vec{r}}}{(2\pi\hbar)^{3/2}}\Phi(\vec{p},t)\, d^3p.\end{aligned} \tag{2.132}$$

Let us use this result to find out the shape of the operator of place-vector in momentum-representation:

$$\begin{aligned}\hat{\vec{r}}\,\Psi(\vec{r},t) &= \int \vec{r}\,\frac{e^{(i/\hbar)\vec{p}\cdot\vec{r}}}{(2\pi\hbar)^{3/2}}\Phi(\vec{p},t)\,d^3p \\ &= \int \left[\frac{\hbar}{i}\frac{\partial}{\partial\vec{p}}\frac{e^{(i/\hbar)\vec{p}\cdot\vec{r}}}{(2\pi\hbar)^{3/2}}\right]\Phi(\vec{p},t)\,d^3p \\ = \int \frac{e^{(i/\hbar)\vec{p}\cdot\vec{r}}}{(2\pi\hbar)^{3/2}}&\left[-\frac{\hbar}{i}\frac{\partial}{\partial\vec{p}}\Phi(\vec{p},t)\right]d^3p,\end{aligned} \tag{2.133}$$

where we can read the operation of the operator of place-vector in momentum-representation:

$$\hat{\vec{r}} \Rightarrow -\frac{\hbar}{i}\frac{\partial}{\partial\vec{p}} \quad . \tag{2.134}$$

Naturally the commutation relation $[\hat{\vec{p}},\hat{\vec{r}}] = \hbar/i$ is valid for both coordinate- and momentum-representation.

2.6.7 Density matrix

Using the matrix formalism we can alter the formula of the quantum mechanical average (or in other words expectation value) into a new shape. If we use the resolution $|\Psi\rangle = \sum_m c_m|m\rangle$, $\langle\Psi| = \sum_n c_n^*\langle n|$, we can rewrite the formula 2.101 into a new expression:

$$\langle\hat{A}\rangle_\Psi = \sum_m\sum_n c_n^*\langle n|\hat{A}|m\rangle c_m = \sum_m\sum_n \rho_{mn}A_{nm} = Tr(\hat{\rho}\hat{A}), \tag{2.135}$$

where we introduced the

$$\rho_{mn} = c_m c_n^*, \tag{2.136}$$

matrix elements of the *density matrix* and corresponsive *density operator*:

$$\hat{\rho} = |\Psi\rangle\langle\Psi|, \tag{2.137}$$

which is a projection operator, projecting onto the state $|\Psi\rangle$. Its continuous matrix shape which is valid in coordinate-representation can be seen below:

$$\rho(x, x') = \langle x|\Psi\rangle\langle\Psi|x'\rangle = \Psi(x)\Psi^*(x'). \tag{2.138}$$

From the time-dependent Schrödinger equation, it is easy to define the motion equation of the density operator:

$$\dot{\hat{\rho}} = |\dot{\Psi}\rangle\langle\Psi| + |\Psi\rangle\langle\dot{\Psi}| = -\frac{i}{\hbar}\hat{H}|\Psi\rangle\langle\Psi| + \frac{i}{\hbar}|\Psi\rangle\langle\Psi|\hat{H}, \tag{2.139}$$

that is

$$\dot{\hat{\rho}} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}]. \tag{2.140}$$

The idea of the density matrix was discovered and introduced by Landau and János Neumann independently, but the motion equation comes from Neumann, hence it is called *von Neumann equation*. The novelty in the density matrix is that the factors $|\Psi\rangle$ and $\langle\Psi|$ of the quantum mechanical average is taken into one object, which can be used in not only the case of a *pure state* $|\Psi\rangle$, but we can average over the ensemble of quantum systems with different quantum states too (*mixed state*). Due to its linearity, Neumann equation remains unchanged after this operation. The connection which we know from 2.135

$$\langle\hat{A}\rangle_\Psi = Tr(\hat{\rho}\hat{A}) \tag{2.141}$$

is valid even in a case like this, but in this situation it generally means double average successively: the quantum mechanical average of a measurable physical quantity in a given quantum state, then a second average over the ensemble of different quantum states.

This ensemble can be imagined as we had many of this quantum system waiting for us to make a measurement on them or we it can mean the ensemble of measurements on a system, which is prepared the same way after each measurement.

Nowadays we have the possibility to make repeated esperiments on single ions or atoms which are trapped in an ion or atom trap.

In the language of the density matrix, the normalization of the wave function means that

$$Tr\,\hat{\rho} = 1, \tag{2.142}$$

that is the average value of the unit operator is 1. This remains unchanged even in the case of mixed state (ensemble)[11]. Since expression 2.137 relating to pure state describes a projector, in case of a pure state also the following formula is valid:

$$Tr\,\hat{\rho}^2 = 1. \tag{2.143}$$

However, if the density operator is averaged over different states, this connection is not true: Generally, for a mixed sate:

$$Tr\,\hat{\rho}^2 \leq 1. \tag{2.144}$$

There is another, very important forma of the density operator: it describes the situations where we want to follow the dynamics of a subsystem which is attached to a larger environment (in Feynman's words: *"the rest of the world"*). The basic phenomenon is the progressive blur of the quantum coherence of the subsystem due to the attaching to the environment (or due to the environmental noise). In a case like this a density matrix takes shape which has a form just like the mixed state density matrix.

2.7 Some of the direct consequences

From the things treated in the last section some general and important results follow, which are the emblematic features of quantum mechanics.

[11] But this is the only property which is insensitive of mixing.

2.7.1 Uncertainty relations

Knowing the recipe of the quantum mechanical average and the idea of self-adjointness, we are ready to derive Heisenberg's uncertainty principles in a mathematically exact way. Before doing it, we still need to know a mathematical tool, namely the *Schwartz inequality* relating to the scalar product, which is the Hilbert space analogue of the inequality $|\vec{a} \cdot \vec{b}| = a\, b\, \cos\alpha \leq a\, b$ relating to vectors:

$$|\langle a|b\rangle|^2 \leq \langle a|a\rangle\, \langle b|b\rangle. \tag{2.145}$$

Let us consider the following self-adjoint operators: $\hat{\Delta A} =: \hat{A} - \langle \hat{A}\rangle$, $\hat{\Delta B} =: \hat{B} - \langle \hat{B}\rangle$. Their commutator is $[\hat{\Delta A}, \hat{\Delta B}] = [\hat{A}, \hat{B}]$. In a state $|\Psi\rangle$, let us apply Schwartz inequality to vectors $|\hat{\Delta A}\,\Psi\rangle$ and $|\hat{\Delta B}\,\Psi\rangle$:

$$\langle\, \hat{\Delta A}\,\Psi \,|\, \hat{\Delta A}\,\Psi\,\rangle\, \langle\, \hat{\Delta B}\,\Psi \,|\, \hat{\Delta B}\,\Psi\,\rangle \geq |\langle\, \hat{\Delta A}\,\Psi \,|\, \hat{\Delta B}\,\Psi\,\rangle|^2. \tag{2.146}$$

Now, we can use the self-adjointness of the operators:

$$\begin{aligned} \langle\, \Psi \,|\, (\hat{\Delta A})^2 \,|\, \Psi\,\rangle\, \langle\, \Psi \,|\, (\hat{\Delta B})^2 \,|\, \Psi\,\rangle &\geq |\langle\, \Psi \,|\, \hat{\Delta A}\,\hat{\Delta B} \,|\, \Psi\,\rangle|^2 \\ = \left|\frac{1}{2}\langle\, \Psi \,|\, \hat{\Delta A}\,\hat{\Delta B} - \hat{\Delta B}\,\hat{\Delta A} \,|\, \Psi\,\rangle + \frac{1}{2}\langle\, \Psi \,|\, \hat{\Delta A}\,\hat{\Delta B} + \hat{\Delta B}\,\hat{\Delta A} \,|\, \Psi\,\rangle\right|^2 &= \frac{1}{4}\big|\langle\, [\hat{A}, \hat{B}]\,\rangle\big|^2 + \frac{1}{4}\big|\langle\, \Psi \,|\, \hat{\Delta A}\,\hat{\Delta B} + \hat{\Delta B}\,\hat{\Delta A} \,|\, \Psi\,\rangle\big|^2, \end{aligned} \tag{2.147}$$

where in the last step we used that the first member of the second row is imaginary, the second one is real, hence their squares are summed. However the second member of the last row is negative, thus we get the following inequality:

$$\langle\, \Psi \,|\, (\hat{\Delta A})^2 \,|\, \Psi\,\rangle\, \langle\, \Psi \,|\, (\hat{\Delta B})^2 \,|\, \Psi\,\rangle \geq \frac{1}{4}\big|\langle\, [\hat{A}, \hat{B}]\,\rangle\big|^2. \tag{2.148}$$

In the end, introducing the notation of $\Delta A =: \sqrt{\langle\, \Psi \,|\, (\hat{\Delta A})^2 \,|\, \Psi\,\rangle}$ we get the customar shape of the uncertainty relation:

$$\Delta A\, \Delta B \geq \frac{1}{2}\big|\langle\, [\hat{A}, \hat{B}]\,\rangle\big|. \tag{2.149}$$

The special case of this is the well known relation: $\Delta\, x\, \Delta\, p \geq \hbar/2$.

It is very important to say again: uncertainty relation between two measurable physical quantities (say A and B) means that the two quantities can not be measured with arbitrary precision simultaneously. Naturally, either of the can be measurable with arbitrary precision but after this measurement the other one will become completely undefined.

2.7.2 The time-derivative of the average value; constants of motion; Ehrenfest's theorem

The average value $\langle\hat{A}\rangle_\Psi = \langle\Psi|\hat{A}|\Psi\rangle$ of an operator of a physical quantity $\hat{A}$ can change in time for two reasons: one of these reasons is the change of the quantum state $|\Psi(t)\rangle$ described by the Schrödinger equation, on the other hand also the physical quantity itself can change: $\dot{\hat{A}} = \partial\hat{A}/\partial t \neq 0$. These changes appear in the time derivation of the average:

$$\begin{aligned} \frac{d}{dt}\langle\hat{A}\rangle &= \langle\dot{\Psi}|\hat{A}|\Psi\rangle + \langle\Psi|\dot{\hat{A}}|\Psi\rangle + \langle\Psi|\hat{A}|\dot{\Psi}\rangle \\ = \frac{i}{\hbar}\langle\Psi|\hat{H}\,\hat{A}|\Psi\rangle &+ \langle\Psi|\frac{\partial\hat{A}}{\partial t}|\Psi\rangle - \frac{i}{\hbar}\langle\Psi|\hat{A}\,\hat{H}|\Psi\rangle \\ &= \frac{i}{\hbar}\langle\Psi|\,[\hat{H}\,,\,\hat{A}]\,|\Psi\rangle + \langle\Psi|\frac{\partial\hat{A}}{\partial t}|\Psi\rangle, \end{aligned} \tag{2.150}$$

where we used the Schrödinger equations 2.111 and 2.114 describing the time dependence of vectors ket and bra. The result in 2.150 is often used in the case when operator $\hat{A}$ can be commuted with the Hamiltonian operator of the system and is not the explicit function of the time:

$$\begin{aligned} [\hat{H},\hat{A}] &= 0, \\ \partial\hat{A}/\partial t &= 0. \end{aligned} \tag{2.151}$$

In this situation its average does not change in time, that is $\hat{A}$ is a *constant of motion*. In addition to this, under these conditions, the total statistics of the measurement of operator $\hat{A}$ is time-independent.

Since the constants of motion can be commuted with $\hat{H}$, they have a common function system with $\hat{H}$.

Let us apply equation 2.150 to analyze the average time dependence of the coordinate and momentum of a single particle: For the sake of simplicity, we can stay at a one dimensional model (w.l.o.g.). Let us consider the well known shape of the Hamiltonian operator $\hat{H} = (\hat{p}^2/2m + V(\hat{x}))$. The results:

$$\begin{aligned} \frac{d}{dt}\langle\hat{p}\rangle = \frac{i}{\hbar}\langle[V(\hat{x}), p]\rangle &= -\frac{i}{\hbar}\left\langle\left[\frac{\hbar}{i}\partial_x, V(x)\right]\right\rangle \\ &= \left\langle -\frac{\partial V(x)}{\partial x}\right\rangle = \langle F(x)\rangle, \end{aligned} \tag{2.152}$$

$$\frac{d}{dt}\langle\hat{x}\rangle = \frac{i}{2m\hbar}\langle[\hat{p}^2, \hat{x}]\rangle = \left\langle\frac{\hat{p}}{m}\right\rangle. \tag{2.153}$$

In these deductions two commutators were used: in the first equation we already knew the used commutator, namely: $\hat{p}_x = (\hbar/i)\partial_x$, but in the second case we needed to derive the following result: $[\hat{p}^2, \hat{x}] = \hat{p}_x[\hat{p}_x, \hat{x}] + [\hat{p}_x, \hat{x}]\hat{p}_x = (2\hbar/i)\hat{p}_x$.

Equations 2.152 and 2.153 are called *Ehrenfest's theorem*. Their content is that the Newtonian motion equations, in a well defined *average* sense, are valid in quantum mechanics: the quantum mechanical average of the momentum defines the average speed of the centre of mass, and the quantum mechanical average of the *force* $F(x) = -\partial V(x)/\partial x$ equals the time-derivative of the average of momentum. These important and nontrivial results can seek a good first insight about the nature of complicated time dependent problems.

2.7.3 Time evolution in Schrödinger and Heisenberg picture

The time evolution of the micro systems obeying the laws of quantum mechanics can be followed via statistical averages with measurement: the system has to be *prepared* in the same initial state many times and after a time t some physical quantity $\hat{A}$ has to be measured. From the complete record of the measured results the average of the functions of $\hat{A}$ – for example the scatter in the uncertainty relations – can be evaluated.

From theoretical aspect the average of the results is given by the equation 2.150 which contains all the information. The start point of the deduction of this equa-

tion was the Schrödinger equation which describes the time evolution of the state vector (*Schrödinger picture*), however the consequence seems to be emphasizing rather the active role of the properties of the averaged operator.

We have already discussed that the time evolution of the state vector is considered a *unitary transformation*, because during its evolution, the norm of the state vector (the scalar product with itself) remains unchanged. Thus it can be written in the following form:

$$|\Psi(t)\rangle = \hat{U}(t)\,|\Psi(0)\rangle, \tag{2.154}$$

where $t = 0$ is an arbitrary chosen initial moment, and $\hat{U}(t)$ is a unitary operator ($\hat{U}\hat{U}^\dagger = \hat{I}$), which satisfies the following dynamical equation, which is a direct consequence of the Schrödinger equation:

$$\partial_t\,\hat{U}(t) = -\frac{i}{\hbar}\hat{H}(t)\,\hat{U}(t), \tag{2.155}$$

where the initial condition is $\hat{U}(0) = \hat{I}$. Let us mark its adjoint equation too:

$$\partial_t\,\hat{U}^\dagger(t) = \frac{i}{\hbar}\hat{U}^\dagger(t)\,\hat{H}(t). \tag{2.156}$$

Using the things treated so far, the average value can be written in the following way:

$$\langle\Psi(t)|\,\hat{A}\,|\Psi(t)\rangle = \langle\Psi(0)|\,\hat{U}^\dagger(t)\,\hat{A}\,\hat{U}(t)\,|\Psi(0)\rangle. \tag{2.157}$$

Here comes the interesting step: let us realize that transition from the first shape into the sacond one is considered as an inverse unitary transformation $\hat{S}(t)$, whose operator is the inverse of $\hat{U}(t)$:

$$\hat{S}(t) := \hat{U}^{-1}(t) = \hat{U}^\dagger(t). \tag{2.158}$$

Let the transformed state vector and operators be denoted by an index "H", referreing to the name of Heisenberg:

$$|\Psi_H\rangle = \hat{S}(t)|\Psi(t)\rangle = \hat{U}^{-1}(t)\hat{U}(t)|\Psi(0)\rangle = |\Psi(0)\rangle; \\ \hat{A}_H(t) = \hat{S}(t)\,\hat{A}\,\hat{S}^{\dagger}(t) = \hat{U}^{\dagger}(t)\,\hat{A}\,\hat{U}(t). \tag{2.159}$$

With this, we have arrived at the *Heisenberg picture*, in which state vector do not change in time; time dependence is carried by the operators. However the time dependence of the average value is unchanged: from equation 2.157, it follows:

$$\langle\Psi(t)|\,\hat{A}\,|\Psi(t)\rangle = \langle\Psi_H|\,\hat{A}_H(t)\,|\Psi_H\rangle. \tag{2.160}$$

For the sake of the completness, we have to give the operator equation which describes the time dependence of the Heisenberg operator $hatA_H(t)$. Using equations 2.154 and 2.155:

$$\frac{d}{dt}\hat{A}_H(t) = \frac{i}{\hbar}[\hat{H}_H(t), \hat{A}_H(t)] + \partial_t\hat{A}_H(t), \tag{2.161}$$

where the last member is responsible for the *explicit* time dependence of the operator $\hat{A}_H\left(\hat{\vec{r}}(t), \hat{\vec{p}}(t), t\right)$.

As we expected the shape of the operator motion equation follows the equation 2.150, which describes the time evolution of the average value with its all consequences. First of all, if there is no time dependence, then the constants of motion are selected by the condition of $[\hat{H}, \hat{A}] = 0$. It is not trivial that Ehrenfest's theorems are valid (in an operator shape) even in Heisenberg picture.

As a really important example, continuity equation 2.53 and current density 2.54 appear in the Heisenberg picture. For a system consisting of particles with place vector operator $\hat{\vec{r}}_i(t)$ and momentum operator $\hat{\vec{p}}_j(t)$, Heisenberg operator of the particle density on a place $\vec{r}$ is evidently

$$\hat{\rho}_H(\hat{\vec{r}}, t) = \sum_i \delta(\hat{\vec{r}} - \hat{\vec{r}}_i(t)). \tag{2.162}$$

If the Hamiltonian operator has the following shape

$$\hat{H}_H(t) = \frac{1}{2m}\sum_i \hat{\vec{p}}_i(t) + V\left(\{\hat{\vec{r}}_i(t)\}\right), \tag{2.163}$$

the Heisenberg-time-derivative of 2.162 can be written into the following informative shape:

$$\frac{d}{dt}\hat{\rho}_H(\hat{\vec{r}},t) = -\vec{\nabla}\cdot\vec{j}_H(\hat{\vec{r}},t), \tag{2.164}$$

where the Heisenberg operator of the current density on a place $\vec{r}$ is

$$\hat{\vec{j}}_H(\vec{r},t) = \frac{1}{2M}\sum_i \Big(\hat{\vec{p}}_i(t)\delta(\hat{\vec{r}}-\hat{\vec{r}}_i(t)) + \delta(\hat{\vec{r}}-\hat{\vec{r}}_i(t))\hat{\vec{p}}_i(t)\Big). \tag{2.165}$$

Besides these pictures, the *interaction picture* is another important description in quantum mechanics but its discussion points beyond the framework of our indtroductory course. For similar reasons, we set aside from the presentation of the *connection between the symmetries of the Hamiltonian operator and the conservation laws of quantum mechanics*, etc.

2.8 The harmonic oscillator: details

In this section we return to a specific system which we have already met. The harmonic oscillator is the most important *solvable*[12] model of the physics describing many of the systems in the nature: the vibrations of molecules and cristals and – after some simple variable replacing – even the vibrations of the electromagnetic field. In addition to this, numerous more complicated physical systems can be modelled if they are described with a model consisting of many oscillators or there are a lot of systems which can be understood if they are considered as harmonic oscillators which are altered (perturbed) a bit. Using the stronger performance of the general formalism, in this section we analize the model of the harmonic oscillator deeper than we did so far.

2.8.1 The algebraic method: creation and annihilation operators and their matrix elements

Using the commutation relations (in case of an oscillator: $[\hat{p},\hat{x}] = \hbar/i$), there are numerous situations when the spectrum of the Hamiltonian operator can be de-

[12] This model can be described completely by analytical (not numerical) calculation.

fined without solving the Schrödinger equation. Let us consider the Hamiltonian operator of the harmonic oscillator:

$$\begin{aligned}\hat{H} = \frac{m\omega^2}{2}\hat{x}^2 + \frac{\hat{p}^2}{2m} = \hbar\omega\frac{m\omega}{2\hbar}\left(\hat{x}^2 + \frac{1}{m^2\omega^2}\hat{p}^2\right)\\ = \hbar\omega\left(\sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - \frac{i}{m\omega}\hat{p}\right)\sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i}{m\omega}\hat{p}\right) + \frac{1}{2}\right)\\ = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right),\end{aligned} \tag{2.166}$$

where in the second row the sum of the square has been rewritten into a product (based on the connection according to which $a^2+b^2 = (a+ib)\times(a-ib)$) but we were careful of the fact that operators $\hat{x}$ and $\hat{p}$ can not be commuted, in the end, in the last row we introduced the following operators:

$$\begin{aligned}\hat{a} &= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i}{m\omega}\hat{p}\right) = \frac{1}{2\sigma}\hat{x} + \frac{i}{\hbar}\sigma\hat{p},\\ \hat{a}^\dagger &= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - \frac{i}{m\omega}\hat{p}\right) = \frac{1}{2\sigma}\hat{x} - \frac{i}{\hbar}\sigma\hat{p},\end{aligned} \tag{2.167}$$

whose commutation relation can be directly calculated from the known commutator of $\hat{p}$ and $\hat{x}$:

$$[\hat{a}, \hat{a}^\dagger] = 1. \tag{2.168}$$

In the definition 2.167, $\sigma = \sqrt{\hbar/(2m\omega)}$, in line with equation 2.82. Operators $\hat{a}$ and $\hat{a}^\dagger$ are the adjoint operators of each other, because $\hat{x}$ and $\hat{p}$ are self-adjoint operators. The name of operator $\hat{a}$ is *annihilation operator* and the name of operator $\hat{a}^\dagger$ is *creation operator*. The thing which is annihilated or created by them is one of the excitation quantums of the oscillator. Let us mark the inverse of the connections 2.167:

$$\hat{x} = \sigma(\hat{a} + \hat{a}^\dagger); \quad \hat{p} = \frac{\hbar}{2i\sigma}(\hat{a} - \hat{a}^\dagger). \tag{2.169}$$

Let us write the Hamiltonian operator into the following shape:

$$\hat{H} = \hbar\omega\left(\hat{N} + \frac{1}{2}\right); \quad \hat{N} = \hat{a}^\dagger\hat{a}. \tag{2.170}$$

Let us focus on the eigenvalues of operator $\hat{N}$ and their normalized eigenvectors:

$$\hat{N}\,|n\rangle = n\,|n\rangle; \quad \langle n|n\rangle = 1. \tag{2.171}$$

Using the commutation relation in 2.168, it is easy to show that besides $|n\rangle$, $\hat{a}|n\rangle$ and $\hat{a}^\dagger|n\rangle$ are eigenvectors too, except when one of them equals 0:

$$\begin{aligned}\hat{N}\,\hat{a}|n\rangle &= (\hat{a}^\dagger\hat{a})\hat{a}|n\rangle = (\hat{a}\hat{a}^\dagger - 1)\hat{a}|n\rangle = (\hat{a}\hat{N} - \hat{a})|n\rangle = (n-1)\,\hat{a}|n\rangle;\\ \hat{N}\,\hat{a}^\dagger|n\rangle &= \hat{a}^\dagger(\hat{a}\hat{a}^\dagger)|n\rangle = \hat{a}^\dagger(\hat{a}^\dagger\hat{a} + 1)|n\rangle = (\hat{a}^\dagger\hat{N} + \hat{a}^\dagger)|n\rangle = (n+1)\,\hat{a}^\dagger|n\rangle.\end{aligned} \tag{2.172}$$

These eigenvectors are not normalized. Using that

$$\|\hat{a}|n\rangle\| = \langle n|\hat{a}^\dagger\,\hat{a}|n\rangle = n, \tag{2.173}$$

we obtain

$$\hat{a}|n\rangle = \sqrt{n}\,|n-1\rangle. \tag{2.174}$$

Similarly, using that

$$\|\hat{a}^\dagger|n\rangle\| = \langle n|\hat{a}\,\hat{a}^\dagger|n\rangle = \langle n|\hat{a}^\dagger\hat{a} + 1|n\rangle = n+1, \tag{2.175}$$

we get

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle. \tag{2.176}$$

According to the equation 2.174, if we start from an eigenvalue n, we can generate a series of eigenstates with eigenvalues $n-1$, $n-2$, etc. However, according to 2.173 no members of this series can have a negative value. This is true, if the

initial value of n is an *integer* number. In this case – through integer values – we get to 0. At this point, the series is interrupted, because according to 2.174:

$$\hat{a}|0\rangle = 0, \tag{2.177}$$

that is no state $|-1\rangle$ is generated, $|0\rangle$ is the ground state of the oscillator. From this ground state all the excited states can be generated by the repeated use of 2.176:

$$|n\rangle = \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n|0\rangle. \tag{2.178}$$

Writing the coordinate representation of $\hat{x}$ and $\hat{p}$ into the shape of operator $\hat{a}$ in 2.167, and using the shape of the ground state wave function in 2.81, we obtain the shape of the normalized wave functions of the excited states, which is appropriate for making calculations:

$$\phi_n(x) = \langle x|n\rangle = \frac{1}{\sqrt{n!\sqrt{2\pi}\sigma}}\left(\frac{x}{2\sigma} + \sigma\frac{d}{dx}\right)^n e^{-\frac{x^2}{4\sigma^2}}. \tag{2.179}$$

With the help of these formulae everything can be calculated about the stationary states of the oscillator. However, it so happened later that nonstationary states obtained interest.

2.8.2 Coherent states (quantum swing)

In 1928, soon after the creation of his wave equation and finding its solution (which describes the harmonic vibrational motion), Schrödinger realized that the highly excited states of the quantum oscillator are *not* the things which correspond to the classical oscillator. In justice, a ground state wave packet seesawing back and forth is the thing that corresponds to the classical oscillator. The shape of this wave packet does not change in time, it does not feather. This interesting solution of the time dependent Schrödinger equation is called coherent state and it has a special importance in the quantum theory of electromagnetic waves, because our world is full of the coherent states of these wave modes; they activate our radios, televisions, mobile phones, etc. As we will see shortly, the coherent state is the superposition of vasty different excited states, with defined factors.

Now we show that the coherent state is the eigenstate of the annihilation operator $\hat{a}$ with the corresponsive eigenvalue. First of all, the normalized eigenvectors $|\alpha\rangle$, which satisfy the eigenvalue equation

$$\hat{a}|\alpha\rangle = \alpha\,|\alpha\rangle, \tag{2.180}$$

can be built from the eigenvectors of the harmonic oscillator:

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}\,|n\rangle, \tag{2.181}$$

which can be verified by direct replacing into equation 2.174. In the next step, we show that this formula – with the corresponsive time dependent eigenvalue α – satisfies the time dependent Schrödinger equation. Let start from the initial state $|\alpha_0\rangle$ and write the known time dependence of the oscillator eigenstates $|n\rangle$:

$$e^{-\frac{|\alpha_0|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha_0^n}{\sqrt{n!}}\, e^{-i\omega(n+1/2)t}|n\rangle = e^{-i\omega\, t/2}|\alpha_0\, e^{-i\omega t}\rangle, \tag{2.182}$$

that is, we reobtained the coherent state in 2.181 – aside from the trivial phase factor corresponding to the "zero point vibration". The parameter of this state is passing round on the complex plane with a frequency ω:

$$\alpha(t) = \alpha_0\, e^{-i\omega\, t} = |\alpha|\, e^{-i(\omega\, t+\delta)}, \tag{2.183}$$

where the initial phase of the $\alpha(t)$ complex is denoted by δ.

How does the wave packet move meanwhile? Using the results in 2.169 let us calculate the time dependent average of the coordinate operator and the momentum operator:

$$\begin{aligned} \langle x\rangle_t &= \langle\alpha(t)|\,\hat{x}\,|\alpha(t)\rangle \\ &= \sigma\langle\alpha(t)|\,\hat{a}+\hat{a}^\dagger\,|\alpha(t)\rangle = \sigma(\alpha(t)+\alpha^*(t)) \\ &= 2\sigma\,\Re[\alpha(t)] = 2\sigma\,|\alpha|\cos(\omega\, t+\delta), \end{aligned} \tag{2.184}$$

$$\begin{aligned} \langle p\rangle_t &= \langle\alpha(t)|\,\hat{p}\,|\alpha(t)\rangle \\ &= \frac{\hbar}{2i\sigma}\langle\alpha(t)|\,\hat{a}-\hat{a}^\dagger\,|\alpha(t)\rangle = \frac{\hbar}{2i\sigma}(\alpha(t)+\alpha^*(t)) \\ &= \frac{\hbar}{\sigma}\,\Im[\alpha(t)] = -\frac{\hbar}{\sigma}\,|\alpha|\sin(\omega\,t+\delta), \end{aligned} \tag{2.185}$$

that is, on the *phase plane* $(x(t),\, p(t))$ the coherent state passes round along an ellipse, whose size is $2\sigma\,|\alpha|$ in x direction, and $(\hbar/\sigma)\,|\alpha|$ in p direction. Meanwhile the diameter of the state remains unchanged: in x direction $2\,\Delta x = 2\sigma$, and in p direction $2\,\Delta p = \hbar/\sigma$.

If we look at the revolving spot of a coherent state $|\alpha| \gg 1$ on the complex plane from far, the spot seems to be a revolvint dot correspondently to the motion on phase plane of the classical mechanics. However $|\alpha|$ shows the average excitation:

$$\bar{n} := \langle n\rangle_\alpha = \langle\alpha|\,\hat{a}^\dagger\,\hat{a}\,|\alpha\rangle = \alpha^*\alpha\,\langle\alpha|\alpha\rangle = |\alpha|^2, \tag{2.186}$$

that is,

$$|\alpha| \approx \sqrt{\bar{n}}, \tag{2.187}$$

so in this average sense, correspondence principle is true: a big value of $\bar{n} \approx$ classical motion. Analyzing the question more carefully, we can realize that the superposition of many excited coherent states is the essential element of the classicalness. This can be seen most directly from the equations in 2.169: $\hat{x}$ and $\hat{p}$ contain the annihilation and creation operators *linearly*. The average value of annilation and creation operators is 0 in any energy eigenstate, they have transitional matrix elements between different eigenstates only. Hence, if we want to have a pendulum with a classical swing (that is, the average of $\hat{x}$ and $\hat{p}$ can be macroscopic), the pendulum has to be a quantum state in which different energy eigenstates are superpositioned.

One more feature of the coherent state has to be mentioned: if the phase $\varphi = \omega\,t + \delta$ of the parameter $\alpha(t)$ is considered as a clock hand and we read the time t and find out the average energy from the average excitation $\bar{n} = \hbar\bar{E}$, then uncertainty relation $\Delta E\,\Delta t \approx \hbar$ can be used to get a new relation between scatter of the phase and the excitation number:

$$\Delta\varphi\,\Delta n \approx 1, \tag{2.188}$$

that is, the coherent state can have a certain phase carrying coherence *because* its excitation number is uncertain: it is the superposition of vasty excited states $|n\rangle$.

2.9 The spin

Though there are still numerous important stores of learning which would have to be treated in an introductory chapter of a standard course on quantum mechanics, at this point we have to finish our introduction so as to have time enough for quantum information in the semester.

However, there is still a significant thing which remains to be discussed, namely the *spin* or rather the mathematical tools that describe it, because these tools are used to draw several similar systems in quantum information. Before learning about the spin we need to discuss some related things – for instance the *Stern-Gerlach experiment* – but we will mention them merely tangentially. During treating them, there will be no effort to be quantum mechanically rigorous, instead we will try to give an expressive – hence necessarily false – description, building on the materials taught in high-school-chemistry and/or physics. Fortunately, from the viewpoint of the consequence, this fault will not matter.

2.9.1 Magnetic momentum and spin

It is well known in quantum mechanics, that besides the *principal (or first) quantum number* (denoted by n, where the possible values of n are 1, 2, 3, ...) of an atom relating to the energy level of the given atom, there are several other quantum numbers too. One of them is the secondary quantum number denoted by l, which is related to the shape of the electron orbital and for a given n its possible values are $(0,\ 1,\ 2,\ 3,\ ...,\ n-1)$. From the exact solution of the Schrödinger equation of the electron of a Hydrogen atom, it follows that the absolute value of the orbit angular momentum vector of the electron is quantized according to the following connection:

$$|\vec{L}| = \hbar\sqrt{l(l+1)}, \tag{2.189}$$

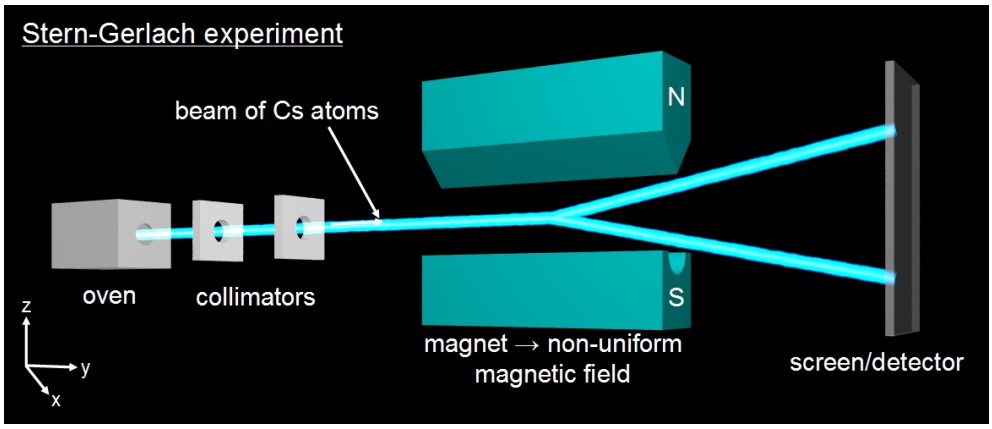


Figure 2.1: The Stern-Gerlach experiment

where $\vec{L}$ denotes the vector of the orbit angular momentum of the electron. The direction of this vector can not be defined precisely. If we select physically a direction (for example using a magnetic field) and make a measurement to define exactly the angular momentum component along the selected direction, directions of other two components will be uncertain. In addition to this, the value of the angular momentum can not change continuously, even in along the chosen direction. For example, considering the selected direction as an x axe of a coordinate system, values of the L_z component of the angular momentum can be the following:

$$L_z = m\hbar, \tag{2.190}$$

where m is the magnetic quantum number and it can have the following values: $m = -l, -l+1, ..., -1, 0, 1, ..., l-1, l$. So the direction of the orbit angular momentum is quantized. This quantization was proved by the *Stern-Gerlach experiment* (1922) (see figure 2.1), where the drift of the atoms was studied in a collimated atomic beam (consisting of silver atoms in the original experiment) led through an inhomogeneous magnetic field which was perpendicular to the direction of the motion of atoms. According to the classical electrodinamics, during its motion a charged particle induces a magnetic field, hence an electron necessarily has a magnetic momentum because of its motion inside the atom. It can be proved that magnetic and angular momentums are commensurable to each other. If it is true in classical physics, it has to be true for also the operators representing physical quantities. The connection between the magnetic and angular momentums can be seen below:

$$\vec{M} = -\frac{e}{2m}\vec{L}, \tag{2.191}$$

where $\vec{M}$ denotes the vector of magnetic momentum. From this expression it follows that in quantum mechanics, the following connection is valid between the operators of the magnetic and angular momentum:

$$\hat{M} = -\frac{e}{2m}\hat{L}. \tag{2.192}$$

From this expression and 2.189, for the quantized magnetic momentum is

$$|\vec{M}| = -\frac{e\hbar}{2m}\sqrt{l(l+1)} = -\mu_B\sqrt{l(l+1)}, \tag{2.193}$$

and it is true for the z component of the magnetic that

$$M_z = -\mu_B m, \tag{2.194}$$

where μ_B is the Bohr magneton. Now we know enough to analyze the Stern-Gerlach experiment. Having no magnetic momentum, the motion of an atom is not disturbed by an external– inhomogenous – magnetic field and the atomic beam remains undiverted. If the atom has a magnetic momentum, but – instead of direction quantization – momentums can trend to arbitrary directions, we would expect the beam to become broader. If the direction quantization "works" and for example $l = 1$, we expect the beam to branch out to three partbeams (in case of $l = 2$, there would be 5 partbeams, etc). The outer electron shell of a Silver atom is just like the outer shell of the Hydrogen atom: there is an electron residing on an s orbit, hence $l = 0$, that is, we do not expect the magnetic momentum to appear. However – back then – experimenters were surprised by the sight of the experimental result, because the initial beam had been split (as it can be seen in figure 2.1) after crossing the inhomogenous magnetic field (that is, there were *two* pieces of partbeams instead of one as it was expected) and the result is obtained if the experiment is made using Hydrogen atoms.

Thus it had to be supposed that electrons have an own magnetic – and angular – momentum which comes from not the "orbital" "motion" of the electrons and its value seems to be $1/2$ (because, in this case momentum can point towards two different directions only, due to $m = \pm 1/2$). This "extra" angular momentum is called *spin*.

It is important to mention that in the Stern-Gerlach experiment, a typical "Neumann measurement" of the magnetic momentum is made (see in 2.6.2) which is closed by the spot on the screen as a detector.

The thing which is split in the Stern-Gerlach experiment is not the mixture but it is the quantum mechanical superposition of the two reverse magnetic momentums: if the inhomogenous magnet is rotated, the two spots on the screen will rotate with the magnet. The theory must describe this strange phenomenon.

Uhlenbeck and and Goudsmit (1925) supposed the magnetic momentum comes from the spinning of the electron[13] considered as an electrically charged pillet. The word *spin* was introduced by them. Hence this phenomenon and the quantum theory of the angular momentum had been irrevocably associated. However, since the work of de Broglie and Schrödinger an electron has been considered as a propagating wave rather than a flying pillet. Hence we need to look for another metaphor, instead of the spinning of a pillet. To think of the circular polarization of a wave is much more better, but it must not be literalized: the "axis of rotation" of the spin is not fixed to the direction of the propagation of the wave. It can trend freely towards arbitrary directions as it was demonstrated by the rotated Stern-Gerlach magnet.

2.9.2 The quantum theory of the half spin: two component spinors and Pauli matrices

In the general quantum theory of the angular momentum (not detailed here), there can be existing angular momentum - eigenstates whose base consists of two orientation; from their complex superposition every other direction can be made. Speaking about spin, let us denote the operator of this kind of angular momentum (which is the internal property of the particle and is not related to any orbital motion) by $\hat{\vec{S}}$. The commutation relations of its vector components (without proving) can be seen below:

$$[\vec{S}_x,\, \vec{S}_y] = i\hbar\vec{S}_z; \quad [\vec{S}_y,\, \vec{S}_z] = i\hbar\vec{S}_x; \quad [\vec{S}_z,\, \vec{S}_x] = i\hbar\vec{S}_y. \tag{2.195}$$

We declare some of the consequences of these relations (without their derivations which can be found in all standard quantum mechanics textbooks):

[13] similarly to the rotation of Earth during it is orbiting around the Sun

- First of all, $\vec{S}_z$ has two eigenvalues: $+\hbar/2$ and $-\hbar/2$.
- The only eigenvalue of $\vec{S}^2$ is $\hbar^2\, s(s+1) = 3/4)\,\hbar^2$, where $s = 1/2$.

The corresponsive eigenvectors have so fundamentally important content that they merit an own, stable notation. Actually – at least – two kinds of symbols are used in Dirac's notation to denote them, and their superpositions are worth writing as a column vector consisting of two components. So the notation kit can be seen below:

$$\begin{aligned} |1/2,\, +1/2\rangle =: |\uparrow\rangle \equiv |+\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |1/2,\, -1/2\rangle =: |\downarrow\rangle \equiv |+\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned} \tag{2.196}$$

The first one is the "spin up", and the second one is the "spin down" – along an axe z. Their superpositions can be seen in the following way:

$$|\chi\rangle = \alpha\,|\uparrow\rangle + \beta\,|\downarrow\rangle \equiv \alpha\,|+\rangle + \beta\,|-\rangle \equiv \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \tag{2.197}$$

where α and β are complex numbers, and due to the normalization

$$|\alpha|^2 + |\beta|^2 = 1 \tag{2.198}$$

is required. We will shortly see that a complex superposition like this describes a spinvector pointing to a definite direction of the three dimensional space. This kind of two component column vector residing in the half spin subspace of the Hilbert space (and which vector transforms in a well defined way during a rotation in the three dimensinal space) was given a specific name: *spinor*; in other words: *Pauli spinor*.

Let us get used to it that the operator of the spinvector is a vector in the three dimensional space and is a 2×2 matrix acting on spinors in the Hilbert space. In the base of the eigenvectors, knowing the eigenvalues, we can write directly that operator $\hat{S}_z$ is drawn by the following 2×2 matrix:

$$\hat{S}_z \Rightarrow \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix}. \tag{2.199}$$

Without proving (which can be found in all standard textbooks on quantum mechanics) we declare the matrices of $\hat{S}_x$ and $\hat{S}_y$:

$$\hat{S}_x \Rightarrow \begin{pmatrix} 0 & \hbar/2 \\ \hbar/2 & 0 \end{pmatrix}; \quad \hat{S}_y \Rightarrow \begin{pmatrix} 0 & -i\hbar/2 \\ i\hbar/2 & 0 \end{pmatrix}. \tag{2.200}$$

Let us introduce the famous *Pauli matrices*:

$$\begin{aligned} \sigma_x &:= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \\ \sigma_z &:= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \hat{I} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{2.201}$$

Using these matrices, the components of spinoperators can be written briefly in the following way:

$$\hat{S}_l \Rightarrow \frac{\hbar}{2}\sigma_l; \quad l = x,\, y,\, z. \tag{2.202}$$

There is a feature of Pauli matrices which can be easily admited, namely: $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \hat{I}$, from this it follows that eigenvalues of σ_x, σ_y and σ_z are ± 1. Due to this, eigenvalues of $\hat{S}_x$, $\hat{S}_y$ and $\hat{S}_z$ are $\pm\hbar/2$. Let us mark the corresponsive eigenvectors too:

$$\begin{aligned} |+\rangle_x &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad |-\rangle_x = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ |+\rangle_y &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} \quad |-\rangle_y = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} \end{aligned} \tag{2.203}$$

The commutation relations of the Pauli matrices are declared in the following way – without dreivation/ proving:

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\,\sigma_l. \tag{2.204}$$

Their anticommutators are perhaps not so trivial:

$$\{\sigma_j, \sigma_k\} \equiv \sigma_j\sigma_k + \sigma_k, \sigma_j = \delta_{jk}. \tag{2.205}$$

As a small practising, let us calculate the components of a spinor belonging to a spin state pointing towards an arbitrary direction given by the polar angles φ and ϑ. This spinor will be the eigenvector (with an eigenvalue $+\hbar/2$) of the spinvector operator projected onto the unit vector $(\sin\vartheta\cos\varphi,\ \sin\vartheta\sin\varphi,\ \cos\vartheta)$ which points towards the given direction. Let us write the matrix shape of the operators:

$$\begin{aligned} \hat{S}_{\varphi,\vartheta} = \sin\vartheta\cos\varphi\,\hat{S}_x + \sin\vartheta\sin\varphi\,\hat{S}_y + \cos\vartheta\,\hat{S}_z = \\ \frac{\hbar}{2}\left(\sin\vartheta\cos\varphi\begin{pmatrix}0&1\\1&0\end{pmatrix} + \sin\vartheta\sin\varphi\begin{pmatrix}0&-i\\i&0\end{pmatrix} + \cos\vartheta\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right) \\ = \frac{\hbar}{2}\begin{pmatrix}\cos\vartheta & \sin\vartheta\,e^{-i\varphi}\\ \sin\vartheta\,e^{i\varphi} & -\cos\vartheta\end{pmatrix}. \end{aligned} \tag{2.206}$$

We need to diagonalize the 2×2 matrix we obtained to get the eigenvectors:

$$|+\rangle_{\varphi,\vartheta} = \begin{pmatrix} e^{-i\varphi/2}\cos\frac{\vartheta}{2} \\ e^{i\varphi/2}\sin\frac{\vartheta}{2}\end{pmatrix};\quad |-\rangle_{\varphi,\vartheta} = \begin{pmatrix} e^{-i\varphi/2}\sin\frac{\vartheta}{2} \\ e^{i\varphi/2}\cos\frac{\vartheta}{2}\end{pmatrix} \tag{2.207}$$

Let us check where they point: $\langle\pm|\hat{S}_x|\pm\rangle = \pm(\hbar/2)\sin\vartheta\cos\varphi$; $\langle\pm|\hat{S}_y|\pm\rangle = \pm(\hbar/2)\sin\vartheta\sin\varphi$, just as we expected.

2.9.3 Spin rotation, quasi-spin, qubit

It is time to ponder over the fact that in the expression 2.207 half angles appear. It means that if any of the angles φ or ϑ are rotated by 2π (a whole rotation), though the average of the spinvector will returns to its initial value, the quantum mechanical state vector will not do the same: it needs a double rotation (with an angle of 4π) to return once.

It is not cheating and not a mistake: with the interference of neutrons whose spin are rotated in a magnetic field, periodicity of 4π can be directly observed. This is an effect of the half spin but in our course we will not discuss it.

Besides describing the electron and other elementary particles, the mathematics of the half spin can be used to describe every *two-state* quantum system. These

systems are often called *quasi-spin* and their dynamics is the time evolution of the superposition of the two states.

These two-state systems can be used to a lot of things, if we can keep up the coherence of the superposition for a time enough. One of the most important application is the information processing. One bit of information can be encoded into each of the two states: one of them is ”yes”, the other is ”no”. Quantum mechanics superadd the superposition to it. Superpositions of ”yes” and ”no”: this is the quantum bit, or in shortly: *qubit*. Possibilities of information processing are multiplied by the superposition: this is the dream of *quantum informatics*. The main block of its fulfilment is the coherence: the environment – as a noise source – washes away the superposition. This happens sooner than a quantum computer can finish the calculation. However, the hoped-for advantages are colossal; the research proceeds intensivly.

To find two-state systems is not an easy task. Atoms and other real systems have a lot of states and from them, experimenter have to be able to select two states somehow. Typical approximations: cooling the system to a very low tempreture to exclude many of the excited states; tuning sharply a very good resonator to the energy difference of two states.

2.10 Afterword for the chapter

In this introductory chapter we tried to seek a summary from the basic knowledges in quantum mechanics to make the understanding of the following chapters (on quantum information) easier. Naturaly, we needed to omit numerous parts of the fundamental knowledge of quantum mechanics because of the character of the course as we have already refered to it several times. Hence – evidently – this textbook can not be considered as a standard course book on quantum mechanics. Fortunately we alreaday know enough to begin to learn about the basic elements of the field of quantum information. However, let us keep in mind that there are still a lot we do not know about the grassroots of quantum mechnics. Nevertheless, there is no need to worry about it, because we will discuss what we have to know when it is necessary.

Chapter 3

The inscape of quantum bits

3.1 Foreword for the following chapters

At the end of the introductory chapter we remarked that there are still a lot we have not learned about the grassroots of quantum mechanics yet, but knowing enough to start to deal with the basic elements of the quantum information, it is time for us to begin to treat the new topic and in case we meet a field where more quantum mechanical knowledge is required than we already have, we will discuss what we need to know. In addition the reader will meet parts where we repeat discussed things explained from another viewpoint – for the sake of a deeper understand.

As for *this* chapter: in the following pages we begin to learn what a quantum bit (or in short: qubit) is. For a deep understanding, first we show a system which can be considered as a qubit. The mentioned system is a neutron which crosses an interferometer (in the famous three-eared interferometer). In this interferometer neutrons show a very surprising behavior that leads us to understand the exact meaning of *Born's* interpretation. Since – as we will see – being a qubit, the behavior of the analyzed system is valid for every system realizing a qubit. So, if we understand this behavior, we will understand the abilities of qubits in general. But, for the present, before starting to explain things written above, we have to say something about the interference itself. Hence the first point of this chapter is the display of *Young's* double slit experiment.

We remark that our explanation in part follows the guideline of *Péter Hraskó*'s[1]

[1]Péter Hraskó taught theoretical physics to the writer of these rows. Gratitude to him, the author of this textbook got a deep insight into the quantum mechanics. Being a successful

which is presented in his very good university lecture notes on quantum mechanics titled *Kvantummechanika*.

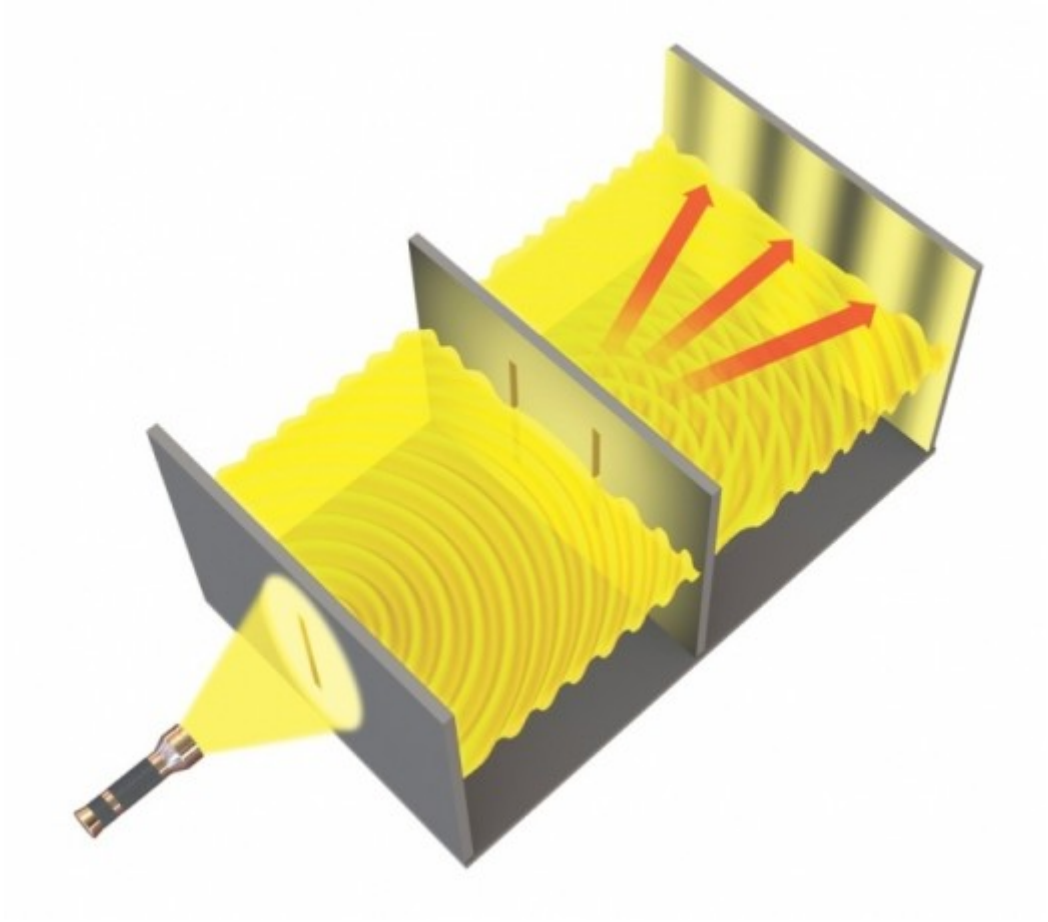
3.2 Optical interference and Young's double slit experiment

In the 18^{th} century, scientists did not manage to decide whether light consists of particles or it is a wave phenomenon. At that time, wave-interpretation won[2], because *interference* of light had been dicovered. By the time 19^{th} century began, it had been known that interference is the sure sign of waves, thus the question - at least at first sight - was answered[3]. Our basic phenomenon – namely, the *Young*-experiment – can be seen below.

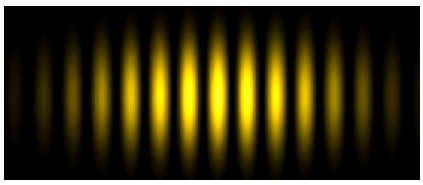
researcher and a very good teacher, his books, notes and papers are worth reading through: `http://peter.hrasko.com/`

[2]Nowadays, it is known that light (and other quantum phenomena) can be featured correctly by wave-particle duality.

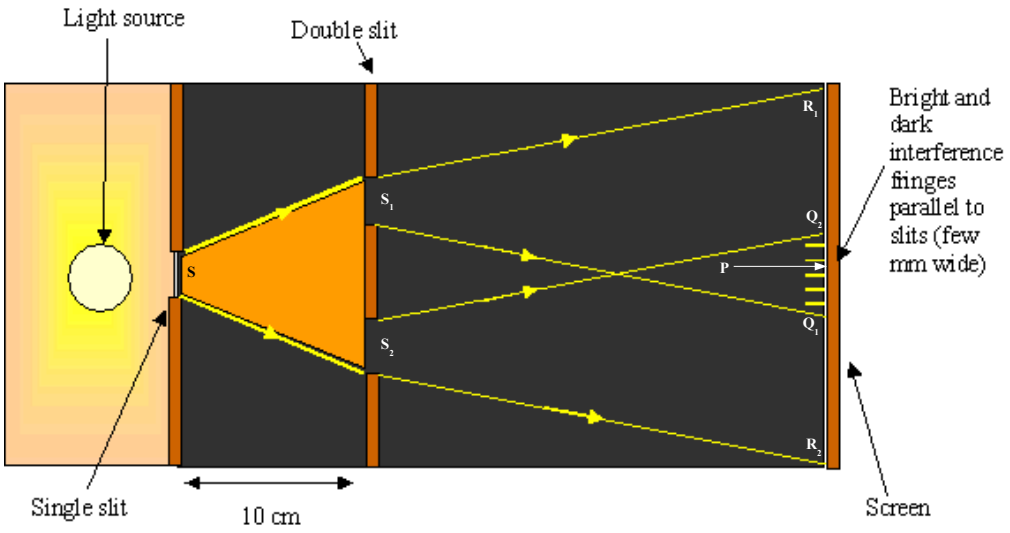
[3]until the birth of quantum mechanics...



As we can see, on the screen interference pattern can be observed in a form like this one, on the right.



Let us consider the figure, below. When slits are narrow, their images expand on the screen Q_1R_1 and Q_2R_2 and overlap each other in the range of Q_1Q_2. In this range (Q_1Q_2), interference pattern comes into being, illuminated and dark lines can be seen.



The *interference-triad*:

- When both slits (S_1 and S_2) are open, in the range of Q_1Q_2 there can be found such a point P which is not illuminated.
- If S_2 is covered (while S_1 is open), in the range of Q_1R_1 illumination will be continuous and point P will be illuminated.
- If S_1 is covered (with an opened S_2), in the range of Q_2R_2 illumination will be continuous and point P will be illuminated again.

Summarized: on the screen there is a point (P) whose illumination *increases* due to *covering* of one of the slits (or, in other words, there is a point (P) whose illumination *decreases* if we *open* both slits).

Let us try to understand this phenomenon by thinking of light as a thing which consists of particles. A particle is localized on its own volume at all times, thus trajectories of particles that go through – for example – slit S_1 can not depend on the state (open/closed) of slit S_2. Some parts of the particles that go through S_1 reach point P and make it illuminated. How can this illumination *break off* due to *opening* of S_2? If the materia which covers S_2 could affect on particles which go through S_1, the observed phenomenon would be explainable. But this explanation is in contradiction to the fact that a materia can not affect on a beam of light which is not in the very near surroundings of the materia. It would be a more

serious attempt, if one tried to understand the interference-triad via considering it as an effect of the particles on each other: when both slits are open, particles of beams have an effect on each other and this effect extrudes light from point P. If this explanation was true, interference-lines would equalize as we use weaker and weaker beam. According to the experiments, the mentioned equalizing never happens: *charasteristic of the interference pattern does not depend on the intensity of beam*. Interference can be observed even in the case when maximum one "particle" is in the apparatus.

On the other hand, turning to the wave-picture to explain the phenomenon, we do not have problems anymore: in case of being both slits open, after starting from S_1 and S_2, part-beams will amplify and degrade each other on the screen, depending on their optical path-difference (phase-difference). Therefore we continue our analysis using the wave-picture.

Conditions of interference in Young-experiment can be summarized in the following four points: 1.diffraction, 2.superposition theorem, 3.monochromacity, 4.coherence. Diffraction means that shadows of objects and images of slits are diffused. Even diffraction itself is a consequence of wave-nature. In an interference-experiment, diffraction is important, because it makes sure the overlaping of part-beams.

Superposition theorem states that superposition of free propagating waves is a possible motion too. Addition of part-beams (according to their signs) in points of the screen is a consequence of this theorem.

Requirements 3 and 4 are necessary to make the value of phase-difference constant, in points of the screen. If this requirement is not satisfied in the experiment, interference pattern will be diffused or it will not come into being. This will happen, if light is not monochromatic, or in other words, its wave-length changes. The same can happen - even in the case of monochromatic beams - if their *initial* phase-difference is not a constant value. This can occur, if slits are illuminated by independent sources of light.

Part-beams are *coherent*, if their phase-difference is constant in time. This can be ensured by creating both beams from the wave-front of the same atom. In Young-experiment common source of light is slit S, which spreads the light coming from left due to diffraction. If apparatus is mechanically stable (optical path-lengths are constant), coherence of slits S_1 and S_2 (as light-sources) can be garanteed by this trick.

It is obvious that there can be degrees between total lack of coherence and total

coherence, thus it is worth speaking about the measure of coherence. This can be represented by sharpness of interference pattern, alias *contrastyness*:

$$K = \frac{I_{max} - I_{min}}{I_{max} + I_{min}},$$

where I_{max} and I_{min} are the maximal and minimal illumination. In an optimal case $K = 1$, or in other words, illumination of dark ranges equals 0, coherence is complete.

3.3 De Broglie-waves

Summarizing what we wrote in this chapter so far, we may think that light is a wave phenomenon and this question is not a question anymore, because interference-triad was demonstrated and from its analysis we know that it can be explained using wave-picture only. Thus interference-triad is the sure sign of wave-nature. But..... – as we mentioned it in the introductory chapter – in 1922, Albert *Einstein* was awarded the Nobel Prize in Physics for his discovery of the law of the photoelectric effect. At this point, for us, its basic statement is important: in his work - in 1905 - *Einstein* supposed that light consists of particles. This was the famous photon-hypothesis. Let us remember the linear harmonic oscillator. In that case, Max *Planck* realized – in 1900 – that oscillator (as a mechanical system) has a quantized motion. *Einstein* generalized this realization to the electromagnetic field. He stated that a light beam (with a pulsatance ω) consists of massless particles (light quanta, or photons) with an energy $\hbar\omega$ or $h\nu$[4]. Though it can seem unbelievable, today it has already known that light can act like particle and it can act like wave too, depending on its circumstances[5]. According to *Einstein*'s reasons, photons are really existent objects. Besides *theory of special relativity*, he gave also other arguments to support his idea: since in *Maxwell*'s electrodynamics a momentum E/c belongs to the energy (E) of electromagnetic waves, a photon (with an energy $h\nu$) also has a momentum $p = h\nu/c = h/\lambda$. Thus photons can be considered as real elmentary particles.

In 1924 Louis *de Broglie* discovered that if - vice versa - a motion of some *matter-wave* can be assigned to an electron which moves with a momentum $p = m_e v$ and

[4]This hypothesis was believed by nobody for almost 20 years, because of the result of the Young-experiment. It seemed to be the resuscitation of *Newton*'s picture of light-particles. Photon-hypothesis became plausible only after creating of modern quantum mechanics.

[5]We will give a more detailed description of this kind of behavior in case of neutrons.

if *Einstein*'s formula

$$\lambda = \frac{h}{p} \tag{3.1}$$

is true for this motion, then *Bohr*'s quantum condition will be given a fantastic simple explanation. Let us look the reason of this statement. *Bohr*'s quantum condition says that there are allowed trajectories[6] of an electron which is inside of an atom. Radii of these allowed trajectories are selected from clasically possible ones by the condition

$$L = n\frac{h}{2\pi} \equiv n\hbar, \tag{3.2}$$

where n is an integer number, L is the angular momentum of electron and $L = pr = m_e vr$. Really, if we rearrange equation 3.2 and we consider *Einstein*'s formula 3.1, we get:

$$2\pi r = n\frac{h}{p} = n\lambda$$

Since n is an integer, this formula means that the wave which "runs" around the nucleus closes into itself. In other words, it can continue its stationary wave-motion, unbroken, perpetually, with a definite (in mathematical language: single-valued) phase. The 3.1 *De Broglie*-relation, which makes a connection between wavelength of an electron moving as a wave and momentum of an electron moving as a ball, is one of the essential connections of quantum mechanics. It is true for not only photons and electrons, but for every quantum mechanical motions. Of course, we have to be careful, because momentum is a vector, not a scalar. Hence we have to introduce the wave-vector $\vec{k}$, which is perpendicular to the wave-front and its size is $k = 2\pi/\lambda$. Using this quantity, 3.1 can be corrected in the following way:

[6] Actually, this condition was a part of the "old" quantum mechanics which was a half-classical theory. It was necessary, because from classical electrodynamics it follows that a charged particle, whose motion has a nonzero acceleration, radiates. According to the half-classical assumptions, electrons revolve around the nucleus, so they have a nonzero acceleration. Hence they radiates. Due to this radiation, their kinetic energy should decrease, thus electrons should fall into the nucleus in a very short time. This phenomenon was not observed, therefore *Bohr* supposed there are trajectories where electrons do not radiate. He called these trajectories *allowed* trajectories. In modern quantum mechanics, there is no motion - in a classical sense. Hence electrons can not have trajectories.

$$\vec{p} = \hbar\vec{k}$$

So, based on *Einstein*'s work, Louis *de Broglie* stated that besides being particles, electrons (and actually all kind of quantum objects) have a wave-nature too[7].

All of these has already been discussed in the introductory chapter, but there can be people who think it is hard to believe. They are right, it is. Hence we have to ask the following questions:

Can it be true? Can something be a particle and a wave too? What does quantum mechanics say about this craziness? Let us turn to the well known *Schrödinger* equation:

$$\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} + \hat{H}\Psi = 0 \tag{3.3}$$

As we know, this equation describes how the physical state of a system changes in time. Now, we are interested in the motion of – for example – a free particle with a mass m. In this case Hamiltoniam operator $\hat{H}$ equals the operator of kinetic energy:

$$\hat{H} = -\frac{\hbar^2}{2m}\triangle \tag{3.4}$$

Hence the equation which describes the physical state of the particle is

$$\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} - \frac{\hbar^2}{2m}\triangle\Psi = 0\,. \tag{3.5}$$

From the structure of this equation it follows that Ψ can be rewritten as a product:

$$\Psi(x, y, z, t) = \psi(x, y, z)T(t), \tag{3.6}$$

where ψ depends on space coordinates only, and T depends on time only. It can be shown that after replacing expression 3.6 into 3.5, we get two equations. One

[7] Actually, *de Broglie* had another, own interpretation to explain it, which differs a bit, but today there are merely a few physicists who share it.

of them contains only the factor that depends on time, and the other one contains the space-depending factor only:

$$\frac{\mathrm{d}T}{\mathrm{d}t} + \frac{i}{\hbar}ET = 0\text{,} \tag{3.7}$$

$$\triangle\psi + \frac{2mE}{\hbar^2}\psi = 0, \tag{3.8}$$

where E is the energy-eigenvalue of the particle. Solution of equation 3.7:

$$T(t) = e^{-\frac{i}{\hbar}Et} \tag{3.9}$$

Let us look for the solution of 3.8 in the shape of

$$\psi = X(x)Y(y)Z(z)\,. \tag{3.10}$$

After replacing this expression into 3.8 and deviding by ψ, we get the following equation:

$$\frac{1}{X}\frac{\mathrm{d}^2\mathrm{X}}{\mathrm{d}x^2} + \frac{1}{Y}\frac{\mathrm{d}^2\mathrm{Y}}{\mathrm{d}y^2} + \frac{1}{Z}\frac{\mathrm{d}^2\mathrm{Z}}{\mathrm{d}z^2} + \frac{2mE}{\hbar^2} = 0 \tag{3.11}$$

This equation is true for every value of the independent variable in only case when

$$\frac{1}{X}\frac{\mathrm{d}^2\mathrm{X}}{\mathrm{d}x^2} = -\alpha^2\text{,}$$

$$\frac{1}{Y}\frac{\mathrm{d}^2\mathrm{Y}}{\mathrm{d}y^2} = -\beta^2\text{,} \tag{3.12}$$

$$\frac{1}{Z}\frac{\mathrm{d}^2\mathrm{Z}}{\mathrm{d}z^2} = -\gamma^2\text{,}$$

where α, β, and γ are constants and they satisfy the connection:

$$\alpha^2 + \beta^2 + \gamma^2 = \frac{2mE}{\hbar^2}\,. \tag{3.13}$$

Solutions of the equations 3.12:

$$X(x) = A_1 e^{i\alpha x}\,,\quad Y(y) = A_2 e^{i\beta y}\,,\quad Z(z) = A_3 e^{i\gamma z}\,. \tag{3.14}$$

Hence, state function of a free particle is

$$\Psi(x,y,z,t) = A e^{i(\alpha x + \beta y + \gamma z - \frac{E}{\hbar}t)}\,. \tag{3.15}$$

Requirement of regularity allows only real values for α, β, and γ. This state function equals the energy-eigenfunction of a free particle, in addition, this is the eigenfunction of components of the momentum, because:

$$\frac{\hbar}{i}\frac{\partial\Psi}{\partial x} = \hbar\alpha\Psi\,,\quad \frac{\hbar}{i}\frac{\partial\Psi}{\partial y} = \hbar\beta\Psi\,,\quad \frac{\hbar}{i}\frac{\partial\Psi}{\partial z} = \hbar\gamma\Psi\,. \tag{3.16}$$

If eigenvalues of components of the momentum $\hbar\alpha$, $\hbar\beta$, and $\hbar\gamma$ are denoted by p_x, p_y, and p_z, state function ca be rewritten into another shape:

$$\Psi(x,y,z,t) = A e^{\frac{i}{\hbar}(p_x x + p_y y + p_z z - Et)}\,. \tag{3.17}$$

According to the classical physics, the connection betwen energy-eigenvalue E and components of the momentum p_x, p_y, p_z has the following form:

$$E = \frac{p_x^2 + p_y^2 + p_z^2}{2m}\,. \tag{3.18}$$

State function 3.17 describes a plane-wave. Expression of a plane-wave which heads towards the direction of the unit-vector $\vec{n}$, and which has a frequency ν and a wavelength λ, can be seen below.

$$\Psi(x,y,z,t) = A e^{2\pi i(\frac{n_x x + n_y y + n_z z}{\lambda} - \nu t)} \tag{3.19}$$

Comparing expression 3.17 to expression 3.19, some connections can be realized between momentum of the particle and wavelength, and between energy of the particle and frequency of the wave:

$$\frac{p_x}{h} = \frac{n_x}{\lambda}, \quad \frac{p_y}{h} = \frac{n_y}{\lambda}, \quad \frac{p_z}{h} = \frac{n_z}{\lambda}, \tag{3.20}$$

$$\frac{E}{h} = \nu. \tag{3.21}$$

Let us square equations 3.20 and add them to each other:

$$\frac{p_x^2 + p_y^2 + p_z^2}{h^2} = \frac{n_x^2 + n_y^2 + n_z^2}{\lambda^2} \tag{3.22}$$

Since $n_x^2 + n_y^2 + n_z^2 = 1$ and $p_x^2 + p_y^2 + p_z^2 = p^2$, from equation 3.22, it follows that

$$\frac{p}{h} = \frac{1}{\lambda}, \quad \text{or} \quad \lambda = \frac{h}{p}. \tag{3.23}$$

So, *according to modern quantum mechanics, the state function of a free particle is a plane-wave* whose frequency can be calculated from the energy of the particle (according to equation 3.21), and its wavelength can be obtained from the momentum of the particle, as we can see in 3.23. Let us remember that between the energy and the feequency of a photon, and between the wavelength and the momentum of a photon, the same connections are valid, and they had already been known by *Einstein*'s work, before quantum mechanics was created.

Giving a probability interpretation of sate function, we required that the volume integral of $|\Psi|^2$ has to be equal to 1. State function 3.17 does not satisfy this requirement or in other words this state can not be normalized. On the other hand, it can be shown that state function of a free particle is actually a wave packet which moves with an average intensity p and its speed equals p/m. The state function of the wave pocket is a normalized and therefore a "legal" state. This wave pocket is called *de Broglie-wave*. Fortunately, solutions of physical problems which are connected to free particles, mostly depend on the value of average momentum p only, thus we can use 3.17 plane-wave to describe the motion of a free particle whose momentum is p.

This plane-wave can also be written in the form below:

$$\Psi(\vec{r},t) = Ae^{\frac{i}{\hbar}(\vec{p}\vec{r}-Et)} \tag{3.24}$$

In literature, mostly $\vec{k} = \vec{p}/\hbar$ is used instead of $\vec{p}$, hence state function which describes the motion of a particle with a momentum p, can be written as

$$\Psi(\vec{r},t) = Ae^{i(\vec{k}_p\vec{r}-\omega_p t)}\,. \tag{3.25}$$

Based on the things that were mentioned here, we can say that the state function of a free particle is a plane-wave whose frequency and wave vector are defined by the formula $\omega = \frac{E}{\hbar}$ and by *de Broglie formula* $\vec{k} = \frac{\vec{p}}{\hbar}$. This function was the first known state function, so this fact explains why state function is called wave function too, generally.

If a particle can be a wave too, it has to show interference phenomenon. Let us see what does our wave function say about it. Let us imagine a plane-wave which is splitted into two parts. After coming paths with different pathlengths, these part-waves will be united. After meeting of part-waves, state can be described by the sum of part-waves:

$$\Psi = \Psi_1 + \Psi_2\,. \tag{3.26}$$

Frequencies (or wavelengths) of the part-waves decribed by Ψ_1+Ψ_2 are the same, but there is a phase-difference between them, due to the difference of the lengths of paths they came. For the sake of simplicity, let us suppose both part-waves along the x-axe of a Cartesian coordinate system, and d denotes the value of the path-difference between them. In this case:

$$\Psi_1 = A_1e^{i(kx-\omega t)}\text{,} \tag{3.27}$$

$$\Psi_2 = A_2e^{i(k(x+d)-\omega t)}\,. \tag{3.28}$$

Summation wave:

$$\Psi = \Psi_1 + \Psi_2 = A_1e^{i(kx-\omega t)} + A_2e^{i(k(x+d)-\omega t)}\,. \tag{3.29}$$

Summation intensity is proportional to $|\Psi|^2$, namely:

$$|\Psi|^2 = |A_1|^2 + |A_2|^2 + A_1^* A_2 e^{ikd} + A_2^* A_1 e^{-ikd}\,. \tag{3.30}$$

Let us rewrite complex numbers A_1 and A_2 into the shape below:

$$A_1 = a_1 e^{i\delta_1}\,, \quad A_2 = a_2 e^{i\delta_2}\,. \tag{3.31}$$

At splitting, phases of the two part-waves are equal to each other, thus $\delta_1 = \delta_2 = \delta$.

$$|\Psi|^2 = |a_1|^2 + |a_2|^2 + 2a_1 a_2 \cos kd, \tag{3.32}$$

where $|a_1|^2$ and $|a_2|^2$ are proportional to intensities of part-waves. Let us suppose that $a_1 = a_2 = a_3$. In this case:

$$|\Psi|^2 = 2a^2(1 + \cos kd)\,. \tag{3.33}$$

If $d = \lambda\frac{2s+1}{4}$, (where s is an integer), then

$$\cos kd = \cos(2s+1)\frac{\pi}{2} = 0\,, \tag{3.34}$$

hence $|\Psi|^2 = 2a_2$, in other words summation intensity equals the sum of part-intensities. If $d = \lambda s$, $|\Psi|^2 = 4a_2$ (summation intensity equals the double of the sum of part-intensities). However, in the other case, when $d = \lambda\frac{2s+1}{2}$, summation intensity $|\Psi|^2 = 0$, two part-waves wipe out each other.

Let us remember, in the case above, we described the interference of not a classical (for example: electromagnetic) wave, but a *de Broglie*-wave which belongs to the motion of a free particle (for example: electron) whose state is described by quantum mechanics. From this, it follows that intensity of the wave (which features the physical state of the particle) is the density of finding-probability of the particle. This means that there will be places in the space where this probability will be zero, and there will be places where this probability is maximal. In other words, distribution of finding-probability of the particle shows an interference pattern.

3.4 Neutrons

In the last section, from a simple solution of *Schrödinger* equation, it follows that state function of a free particle is a plane-wave. When we say the word *particle*, we think of all the objects whose behaviour is described by quantum mechanics (atoms, molecules, subatomic particle, etc). It was a natural intention of experimental physicists to study particles experimentally, whether they really show a wave-nature too, besides being particles. As we have already mentioned, interference is the sure sign of waves, so if interference-triad can be demonstrated, we can say that the studied phenomenon is a wave phenomenon. Our method is to degrade the beam, till its intensity shows that maximum one particle can be ”contained” by the beam. After doing it, knowing that one particle can not be in several separated places which are far from each other and observing interference, we can say that we demonstrated the wave-nature of the given object.

The question is how can we do such a *Young*-type experiment? First of all, we have to look for such kind of particle that can be treated relatively easily in an experiment. Neutron is an ideal candidate, because

- first of all, we know that neutrons *are particles*, because *when we observe them, we always find particles*,
- our neutron sources gives an intensity which ensures that maximum one neutron is in our experimental device,
- being uncharged, it does not ”feel” the effect of charged particles of matter, hence these effects can not disturb our experiment,
- having relatively big mass, relatively less energy is enough to reach the suitable wave-length [8].

The other question is that can *Young*’s setup be used for studying neutrons? Why does this question arise? From wave-theory, we know that from the point of view of perceptibility, distance between slits has to be equal to range of wavelength (in an optimal case). Let us calculate the slit-distance which is necessary for studying neutrons. First we have to calculate *de Broglie* wavelength of neutron:

$$\lambda_n = \frac{h}{p_n} \tag{3.35}$$

[8] We will shortly understand the last statement.

If the value of the temperature is known (for example: $T = 300K$), energy of neutron can be calculated in the following way: $E_n = k_B T$, where k_B is the *Boltzmann*-constant. Energy of neutron can be expressed by another formula too, namely $E_n = \frac{p_n^2}{2m_n}$. We know that $m_n = 1.7 \times 10^{-27}$kg, thus from the two expresses of energy, p can be obtained:

$$p = \sqrt{k_B T 2 m_n} \tag{3.36}$$

After replacing the calculated p into the 3.35 *de Broglie* formula, we get the value of λ:

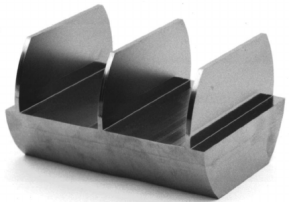
$$\lambda = 1.2 \ \ \mathrm{nm} \tag{3.37}$$

This wavelength is much less than the wavelength of light, and what is more we can not create an experimental setup which has a slit distance of this size. Fortunately, nature gives us a solution for this problem, because distance between atoms in a regular crystal-lattice has the value which is in this range of size. An atom corresponds to a slit, because it spreads *de Broglie*-wave due to diffraction[9] (just as a slit spreads the light beam). Obviously, in crystal-lattice, there are a lot of "slits" (atoms), hence there are a lot of part-beams. In one hand, this is not a problem, because the more slits we have, the bigger value of *contrastyness* (K) we get. On the other hand, this is a problem, because we would like to demostrate the interference-triad. Let us remember what we did before: we had to cover one of the slits (or leave them open). It is clear, we can not cover slits (atoms) in case of crystal-lattice, because they are too many for covering them all.

Fortunately, we are very lucky, because it can be demonstrared by mathematical theory of diffraction that if a beam incidences to the surface of the crystal under an optimal angle (and the thickness of the crystal is also optimal), then the crystal-lattice behaves as a beam splitter (or, in other words, a semi-permeable mirror). This means that it splits the beam into two coherent part-beams[10] and this is exactly what we need, if we would like to achieve all three experiments of interference-triad.

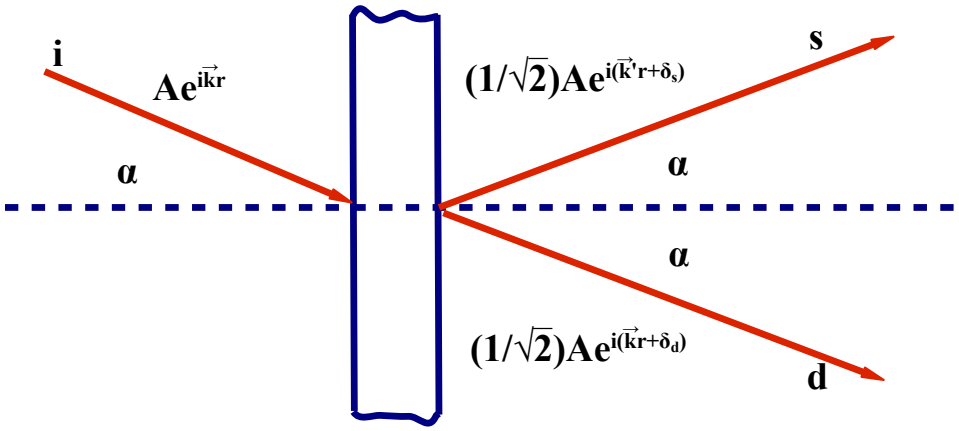
[9]This is possible due to nuclear force that has a very short effective range. This force affects even on the uncharged neutrons.

[10]We will see soon, that while in *Young*-experiment, slits did a wave- front-splitting, in case of crystal-lattice an amplitude-splitting is done.



This effect of crystal-lattice is used in the famous *three-eared interferometer*, which can be seen to the right. This device makes possible for us to achieve *Young*-type experiments with particles. It was developed in the middle of the sixties by *Bonse* and *Hart*. This tool is used in studying neutron beams, since 1974. The basic element of the three-eared interferometer is the "ear". Without any proving, we say that this element works as a beam splitter (it consists of silicon atoms forming a perfect crystal-lattice, and its thickness is also suitable for working as a beam splitter). The "ear" split the incident neutron beam into two coherent part-beams. On the other hand, from the time-mirroring invariance of quantum theory, it follows that part-beams which have a time-mirrored motion will be *united* by the "ear".

Let us study the effect of the "ear" on the wave functions of beams.

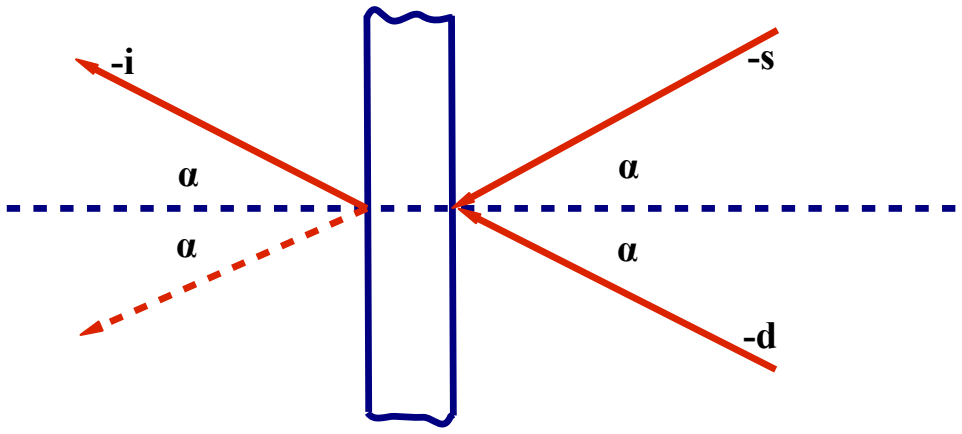


Wave function of neutron which crosses through the metal plate (the ear) is a wave pocket that arrives from direction i (i as incident) and behind the plate, it splits for a superposition of two wave pockets. One of these wave pockets heads along the direction s (as scattered), the other one goes towards the direction d (as direct). For the sake of simplicity, we suppose the metal plate halves the intensity, thus amplitudes of outgoing part-beams equal the amplitude of incoming part devided by $\sqrt{2}$. As we mentioned, in most cases, plane-waves can be used in calculations instead of wave packets. Now, this is exactly a case, like those are. Hence, outside of plate, wave function has the following form:

$$\Psi(x) = \begin{cases} Ae^{i\vec{k}\vec{r}} & \text{, in beam } i \\ \frac{A}{\sqrt{2}}e^{i(\vec{k'}\vec{r}+\delta_s)} & \text{, in beam } s \\ \frac{A}{\sqrt{2}}e^{i(\vec{k}\vec{r}+\delta_d)} & \text{, in beam } d, \end{cases}$$

where common factor $e^{-i\omega t}$ is omitted. Vector $\vec{k}$ shows towards the direction of incident and direct beams, vector $\vec{k'}$ shows towards the direction of scattered beam, and $k = k'$. Phases δ_s and δ_d are created by propagation inside of the plate. So, wave functions of outgoing beams can be obtained from the wave function of incoming beam in the following way:
When we write wave function of scattered beam, after replacing wave vector by the wave vector of the new direction, we have to multiply by the factor $\frac{1}{\sqrt{2}}e^{i\delta_s}$.
In wave function of direct beam, leaving wave vector unchanged, it is enough for us to multiply by the factor $\frac{1}{\sqrt{2}}e^{i\delta_d}$.

Now, we show that due to time-mirroring invariance of quantum theory, difference between the two phases equals $\pi/2$.



Wave function of the time-mirrored motion equals $\Psi^*(x)$ (after omitting the common factor $e^{+i\omega(-t)} = e^{-i\omega t}$). In this case, *two incident* beams reach the plate from right (from directions $-s$ and $-d$). Wave functions of these incident beams are the complex conjugate of parts of $\Psi(x)$ along directions s and d. After crossing through the plate, wave functions can be created from these wave functions in a way which we have already used before: we multiply by the factor $\frac{1}{\sqrt{2}}e^{i\delta_s}$ or $\frac{1}{\sqrt{2}}e^{i\delta_d}$, and in case of change of direction, we replace wave vector by the wave vector of the new direction.

From this rule, it follows that – in the time-mirrored motion – wave function of the

beam which outgoes towards the line of the original motion $(-i)$ has the following form

$$\left[\frac{A}{\sqrt{2}}e^{-i(\vec{k}\vec{r}+\delta_s)}\right]\frac{1}{\sqrt{2}}e^{i\delta_s}+\left[\frac{A}{\sqrt{2}}e^{-i(\vec{k}\vec{r}+\delta_d)}\right]\frac{1}{\sqrt{2}}e^{i\delta_d},$$

provided A is real.This sum equals $Ae^{-i\vec{k}\vec{r}}$, which is the complex conjugate of the incomming (original) plane-wave. As a consequence of time-mirroring invariance, this has to be true. Along the other direction (denoted by dashed arrow), on the left side of the plate, wave function has to be vanished:

$$\left[\frac{A}{\sqrt{2}}e^{-i(\vec{k'}\vec{r}+\delta_d)}\right]\frac{1}{\sqrt{2}}e^{i\delta_s}+\left[\frac{A}{\sqrt{2}}e^{-i(\vec{k'}\vec{r}+\delta_s)}\right]\frac{1}{\sqrt{2}}e^{i\delta_d}=0\,.$$

After simpification, this equation will have the form below:

$$e^{i(\delta_s-\delta_d)}+e^{i(\delta_d-\delta_s)}=0\,.$$

Let us note that on the left side of the equation, the two members are complex conjugates of each other. This means their sum may equal 0 in the only case when they are pure imaginary. What is more, they are phase factors, thus condition of vanishing of their sum is

$$e^{i(\delta_s-\delta_d)}=i \;\; \texttt{or} \;\; -i\,,$$

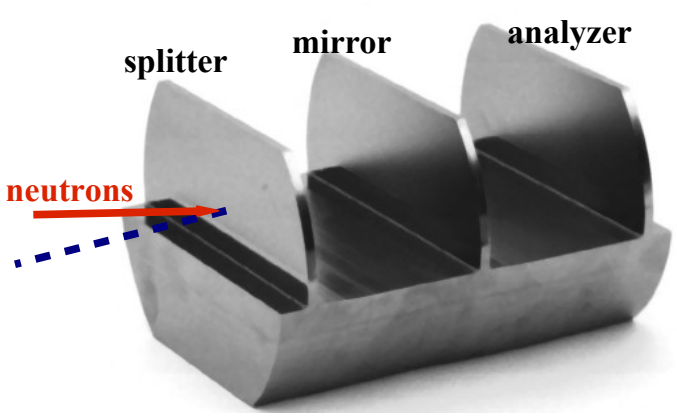
that is $\delta_s-\delta_d=\pm\pi/2$, as we stated. In our case, now, it is all the same which sign we use. From *Schöridinger*-equation, detailed calculation gives the lower one, thus we can choose this one:

$$\delta_s=\delta_d-\frac{\pi}{2}$$

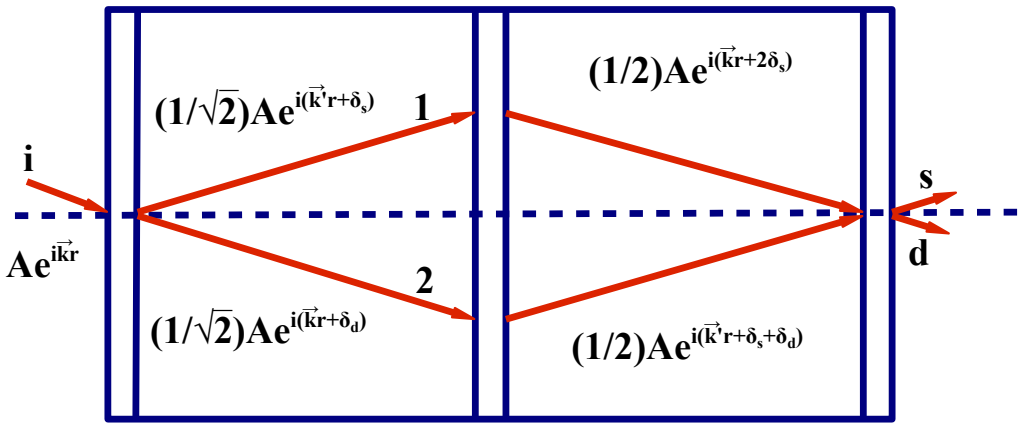
Meaning of this equation can be expressed in the following way: in the original motion, *wave function of scattered beam and wave function of direct beam differ from each other in a factor* $e^{-i\pi/2}=-i$. In respect to time-mirrored motion, *one of the two incoming beams will cange its direction. This will be the one which differs from the other one in a factor* $+i$. Change of sign is a consequance of complex conjugating that is connected to time-mirroring.

3.5 The three-eared interferometer

The essence of the three-eared interferometer is a perfect silicon crystal with a lentgth of $8 - 10$ cm. From this crystal a shape is cut out as it can be seen below.



Left ear is a splitter that splits the beam which comes from a reactor into two part-beams whose intensities equal each other. Middle ear is a *mirror*, which also splits incoming beams into two part-beams, but the device uses only the scattered part-beams. The role of the mirror is to make diverging beams convegrent in order that we can *analyze* their interference. This analysis is achived by the third ear which is the analizator, because interference of part-beams that arrive from different directions can be analyzed by the help of that ear. Below, top-view of the interferometer can be seen, where also wave functions of part-beams are enfaced.



It can be easy realized that motion of neutron to analyzer is the time-mirrored of motion of neutron to the splitter. It is understandable, if we notice that relative

phase of part-beams reaching the analyzer equals $\pi/2$. Really, upper part-beam (let us call it 1) changes its direction several times than beam 2 (down). The difference between their numbers of direction-change equals 1. This means that wave function of beam 1 is the wave function of beam 2 multiplied by $-i$. So, the rule that we found for time-mirrored motion can be applied when part-beams cross through the analyzer. In other words, only one direction will have a nonzero wave function from the two possible directions (s and d) on output of the interferometer. Since wave function of beam 2 equals wave function of beam 1 multiplied by $+i$, beam 2 will change its direction, hence neutrons will appear in direct direction only ($w_d = 1$, $w_s = 0$).

So, we found that neutrons will appear in direct direction on output of the interferometer only. The same result can be obtained by using wave functions enfaced on the scheme without referring to time- mirroring.

But what is in the case of covering one of the inner paths by a cadmium plate which absorbs neutrons? In this case, analyzer becomes a splitter, thus neutrons can be found in both directions (s and d) with equal probability.

Let us notice what it means. *Clear form of interference-triad faces us*: due to *covering* one of the paths, neutrons *appear* in a direction where there were no neutrons before. Covering is not a problem, because distance between part-beams is several centimeters, and we know our reactors ensure automatically that at most one neutron is in the interferometer at a time, so interference can not be caused by their mutual effect on each other.

This experimental result shows *when they are not observed, neutrons do not act like particles*. On the other hand, as we know, when we observe them, we always find them in a well definite point of space, as localized particles. Every micro particles[11] has this "deceitful" property which is called *dual nature* or *wave-particle dualism*. We emphasize, this conclusion was drawn by demonstrating the interference-triad only, quantum theory was not used at all. Theory was needed for designing the device, where Schrödinger-equation of neutrons moving in crystal had to be solved.

[11]This kind of behaviour was observed experimentally even in case of C_{60}, and C_{70} fullerene molecules, by *Zeilinger-Arndt* research group, in 1999. Mass of biggest molecules whose interference were demonstrated succesfully was approximately $3 \times 10^{-24}\ kg$. Molecules with a greater mass are not vaporizable in a stove. Nevertheless, these observations foreshadow that the key which distinguishes quantum objects from their classical mates is not their size.

3.6 From the exact interpretation of Born-hypothesis to quantum bits

Born-hypothesis gives the physical interpretation of mathematical apparatus of quantum mechanics. Specifically, in case of measurement of position, this hypothesis states, after making a position- measurement at a time t, particle whose state function is $\Psi(x, y, z, t)$ will be found in volume $\mathrm{d}V$ around a point with coordinates x, y, z, with a probability $|\Psi(x, y, z, t)|^2\mathrm{d}V$. Let us ask, how much is the probability of observing of the (only) neutron in part-beam 1 (w_1) or in part-beam 2 (w_2) in the interferometer? Let us choose a time t when the entire wave pocket describing the motion of neutron is within the interferometer. Since beam is splitted by the beam splitter, it follows that:

$$w_1 = \int_1 |\Psi(x, y, z, t)|^2\mathrm{d}V = \frac{1}{2}$$

$$w_2 = \int_2 |\Psi(x, y, z, t)|^2\mathrm{d}V = \frac{1}{2},$$

where ranges of integrals cover volumes of part-beam 1 and part-beam 2.
How can we check this result? We can take detectors into the interferometer (into both part-beams). In this case, every single neutron makes only one of the detectors click[12]. On the other hand, in case of observing numerous neutrons, numbers of clicking of detectors will equal each other. In a normal mode, there are no detectors within paths of part-beams. In this case neutrons are monitored by detectors on the output of the interferometer, in directions s and d. But we can ask a quetion which summarizes the essence of basic problem of this section: can we say that when wave pocket is within the interferometer, neutron *locates in one or in the other part-beam with a probability of* $1/2$*, even now?*
That was the case really, when we detected them within the interferometer. But could any change happen due to shifting detectors out of interferometer? People tend to answer this question by saying a categoric "no". This answer implies a picture in which neutrons are within one or in the other part-beam with a probability $1/2$, even now, since neutrons are always observed as integer particles, so part- beam has to be chosen by a neutron at the time of leaving the splitter. It can seem that the way of leaving the splitter can not be changed by inserting detectors into the interferometer (in an arbitrary distance from splitter) or taking out them.

[12]Being an integer particle, neutron can create nuclear reaction which makes detector click, in one of the detectors only.

However, this arguing is denied most definitely by interference of part-beams! As we saw, interference-triad is the surest sign that the thing which crosses through the interferometer (and which is detected as a particle, on the output of the interferometer) can not act like a particle within the interferometer. If this thing would reach the analyzer arriving from one *or* from the other direction, we would not observe any interference.
So, leaving the splitter, neutron chooses either part-beam or the other one in only case, when we can monitor it directly by using detectors taken into both part-beams (within the interferometer) to find out which part-beam has been chosen by the neutron.

Hence the correct answer of our original question is that: when we does not observe the neutron directly, *it is not true* that the neutron *is (resides)* in one or in the other part-beam with a probability $1/2$. Value of probability is correct, this is not the problem. Then.....what's wrong? That picture is false which is implied by our original question: when we do not observe it, neutron does not choose between part-beams: the fact that neutron is within one of the part-beams seems not to exclude that neutron is within the other part-beam too, at the same time.

This conclusion can be expressed by also the following way: *in classical physics, possibilities that are mutually excluded by each other, do not necessarily exclude each other in quantum physics (but they "interfere").* Probably, this statement summarizes the essential difference between classical and quantum physics, in the most briefly way.

Let us notice, contradiction which is represented by this paradox is not a logical one, but (merely?) it enlightens the inadequacy of our paradigm in which phenomena (in this case: the motion itself) are incorporated by us. Let us draw again the statements which are in conflict with each other on the surface.

- When detectors are within the interferometer, every neutron is detected in one or in the other part-beam.
- When detectors are out of interferometer, experimental results do not comport with our *belief* according to which a neutron fares through the interferometer in one or in the other part-beam, even now.

Let us notice, not the same thing is said to be "true" and "false" at the same time, because in the two cases, physical situations are not the same. Nevertheless, we think these two statements paradoxical, because building upon our experiences (or upon our inborn paradigm), we *believe* that choice of beam had to happen when

a neutron left the splitter, and this choice can not depend on further parts of path of neutron. However, this is just an anticipation, an instinctive hypothesis, but it can not be considered a *fact*. Obviously, it is senseless to ask, how much is the probability of observation of neutron by detectors taken into interferometer, when these detectors are not within the interferometer....
This question can be drawn sensibly only in such a way that: how much *would be* the probability of observation of neutron, if detectors *would be* within the interferometer?
In daily life (and also in non-quantum physics), we give bona answers to questions of this kind, without noticing that the answer is *based on an unverifiable hypothesis* according to which validity of the answer does not depend on the alteration of condition. Why would be essential the position of distant detectors (in front of/beyond the analyzer) from the point of view of leaving the splitter? - as we believe instinctively. In case of macroscopic balls, it would not be essential, as we know, since that part of scientific practice which handles macroscopic world is successful, though it is based on this assumption. But when we study neutrons, the mentioned alteration is essential. This is proved by the interference.

As a summary, we can establish that *Born*-hypothesis gives probabilities of results of *real observations* only, and it can not be applied when measurements are merely imagined. It tells not the probability of *being* (or residing) of a particle in a certain position, but the probability of *finding*[13] of a particle on a certain place, in case of making a real measurement (observation) using a suitable device.

3.7 What has this got to do with quantum bits?

What do we know about (classical) bits (in a nutshell)?
They are physical systems that have two states associated with values 0 and 1, and with the help of them, information can be processed and stored. Knowing their behaviour well, we can use them to make our work easier and more effective.

What do we surmise about quantum bits (or in short qubits)?
They are quantum physical systems that have two *observable* (or possible) states (in other words, they are said to be *two-state quantum systems*) associated with values 0 and 1, and with the help of them, (quantum) information can be pro-

[13]Do we remember the footnote that said: "There is a very important reason why we wrote "finding" instead of writing "being". This reason will be explained soon and as we will see, *it will be one of the keys for us to understand how quantum mechanic works*". Well .....the explanation has just been given.

cessed and stored. We can use them to make our work easier and more effective in only case we know their behaviour. (We anticipate in advance that using qubits our possibilities spread in an amazingly large measure.)

Let us notice something. In an interferometer, a neutron represents a *two-state quantum system*, hence it can be considered a qubit. Due to the linearity of dynamical equation, it follows that a superposition of two (or more) states is also a state. It means, if we have two obsevable (eigen)states of a quantum system, any superposition of these states is also a possible state of the system, though it is not an observable state (not an eigenstate). There is no phenomenon like this in classical physics (for example: it would be very strange for me to reside on several places of the space at the same time, wouldn't it be? Nevertheless, quantum objects show this kind of behaviour - as we saw). In case of neutrons, the mentioned observable states were the two part-beams. They were the eigenstates of the system of the neutron. Due to linearity, superposition of them is also a state of the system. (This makes possible for a single neutron to interfere with itself.) Actually, qubits are realized in other ways, but it does not matter, because the same quantum physical description belongs to each system of this kind. Thus, if we understand the behaviour of neutrons in interferometer, we have the basic knowlegde to understand the behaviour of qubits, because there are two observable states in case of qubits too, but qubits (just like neutrons in an interferometer) do not choose from between their observable states. While no measurement is made, a qubit is in the superposition of its observable states, without choosing from between them.

Chapter 4

Qubits

Knowing enough to understand the behaviour of qubits, from this chapter we can begin to learn about elements of quantum information. Naturally first of all the idea of a qubit will be presented as one of the most important actor of this subject. As we will see, a qubit is a state vector in a two dimensional Hilbert-space. Though these ideas have already been introduced in the introductory chapter, it can be useful, if we start this little review by repeating some important knowledges. In a nutshell, from this point we begin to learn about things belonging to the three main points can be seen below:

- storage of information
  - the information is stored in a quantum system
  - information: the state of the quantum system
- processing of information (by using unitary operators)
- readout of information (by making quantum measurement)

4.1 A brief story of qubits

As it is widely known, in classical theory of information, least unit of information is the *bit*, whose value can be either 0 or 1. On the other hand, in quantum information theory, this role is acted by quantum bits, or in short *qubits*. First of all we try to provide a general characterization of qubits. The question naturally arises: what can be qubits made of? The answer: *any two-state quantum system can be considered as a qubit*, where the state of the system is the information itself. At the three-eared interferometer, we saw that state of a quantum system is not necessarily defined. What does it mean, in case of bits? While a classical bit can have merely two possible states (associated with 0 or 1) which are mutually excluded by each other, a qubit can have more than two possible states, because different superpositions of classically possible states (associated with 0 or 1) are also possible states of a qubit (of a quantum system). In a superposition, state of a qubit contains both classically possible (or observable) states with different weights. A qubit which is in the superposition of states titled 0 and 1, till readout of information (or in other words, till making a measurement), does not choose from between 0 and 1. Knowing the state of a qubit, we can predict merely probabilities of "materialization" of states 0 and 1, in case of making a measurement.

Using a quantum computer to process information, our input state can be a superposition of classically possible states. Thus a quantum computer can process different input items in a parallel way, at the same time. This parallelism makes quantum computers able to solve problems which can not be solved by classical computers. (Besides superpositon, there are several other adventages of the use of qubits, but more on this later.) For the present - as we promised - we show what a state vector is.

4.1.1 The state vector

In the general formalism of quantum mechanics (or in other name: *Dirac*-formalism) introduced in the introductory chapter, state of a system is described by a general, infinite dimensional vector. This vector is independent of any coordinate-systems. To understand better the idea of state vectors, let us look at the simplest physical system in which we have a particle with a mass m which moves in a potential $V(q)$, where q is the general coordinate of the particle and it can have any value from $-\infty$ to $+\infty$. As we have already known, in *Schrödinger*-formalism of quantum mechanics (in coordinate representation), at a time t, the state of a particle is described by its wave function $\Psi(q,t)$. On one hand, if no measurement

is made, (from a state $\Psi(q,t_0)$), this state evolves according to the well known *Schrödinger*- equation:

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2}+V(q)\right]\Psi(q,t)=i\hbar\frac{\partial}{\partial t}\Psi(q,t)$$

On the other hand, we know, in case of a measurement, only the probability of obtaining a certain classically possible (or observable, or eigen-) state of the system can be calculated. Let us notice a strange duality:

- if the system is not disturbed by a measurement, its time-evolution is a nice, liquid, and smooth evolution which can be considered as a model of causality.
- in case of inserting a measurement into this smooth time-evolution, the state of the system falls into one of its classically possible states (this phenomenon is called: wave function collapse), and we can calculate merely the value of probability of finding the system in a certain eigenstate.

Staying at this example, after making a position-measurement, we know that the value of probability of finding the particle in the range between q and $q+\mathrm{d}q$ equals $|\Psi(q,t)|^2\mathrm{d}q$.

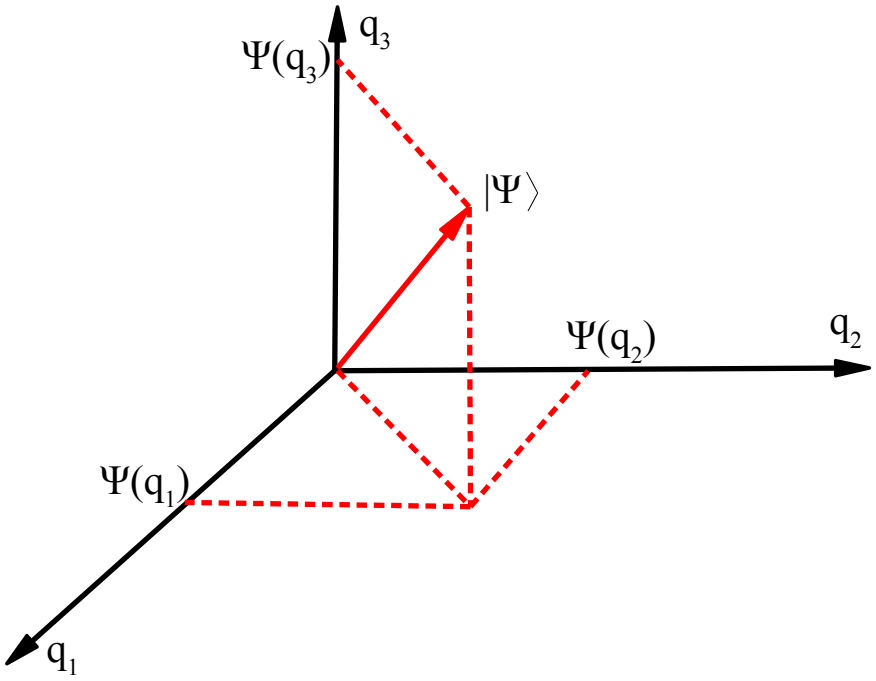
We already know that a state described by $\Psi(q,t)$ can be featured by momentum representation too, where it is desribed by $\Psi(p,t)$. These representations are connected to each other by the *Fourier*-transform:

$$\Psi(p,t)=\frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\Psi(q,t)e^{-\frac{ipq}{\hbar}}\mathrm{d}g$$

Both representations feature *the same* state, and they are totally equivalent to each other. In our case, the role of a represantation is similar to the role of a coordinate system in geometry. Accepting this analogy, the question naturally arises: if problems of geometry can be handled by vectors which are independent of any coordinate system, then similarly, is it possible to describe the state of a quantum system in a way which is independent of coordinate systems (representations)? The known answer is: yes. This is the aim of the *Dirac*-formalism.

Though we have already spoken about this formalism, but it is easily possible that there are several students who did not understand perfectly the explanation. Here we try to provide a more graphic and not-so-dense discussion of this topic.

To introduce this formalism, let us try to give a geometrical interpretation to wave function $\Psi(q,t)$, at a frozen time t. As we mentioned, coordinate q can have any value in the range of $-\infty$ to $+\infty$. On each places q_1, q_2, q_3, etc, wave function has the values $\Psi(q_1)$, $\Psi(q_2)$, $\Psi(q_3)$, etc. According to this conception, we can imagine an infinite dimensional space containing mutually perpendicular axes associated values of q (q_1, q_2, q_3, etc). In this "system of axes", $\Psi(q_1)$ is the projection of some vector on axe q_1, and $\Psi(q_2)$ is the projection of the same vector on axe q_1, and so on. In this case, the mentioned vector (as its components too) represents the state of the system (which was described by state function $\Psi(q,t)$, at a frozen time t), hence this vector is called *state vector*. However, this vector is not an average one, because its components can be even complex numbers, thus there is a special notation for this kind of vectors. The symbol introduced by *Dirac* is the following one: $|\rangle$. A vector whose components are $\Psi(q_1)$, $\Psi(q_2)$, $\Psi(q_3)$, etc is called Ψ *ket* and it is denoted by the symbol $|\Psi\rangle$. On the figure below, we try to illustrate this vector, but unfortunatelly, its visualization can be followed till maximum three mutually perpendicular axes:



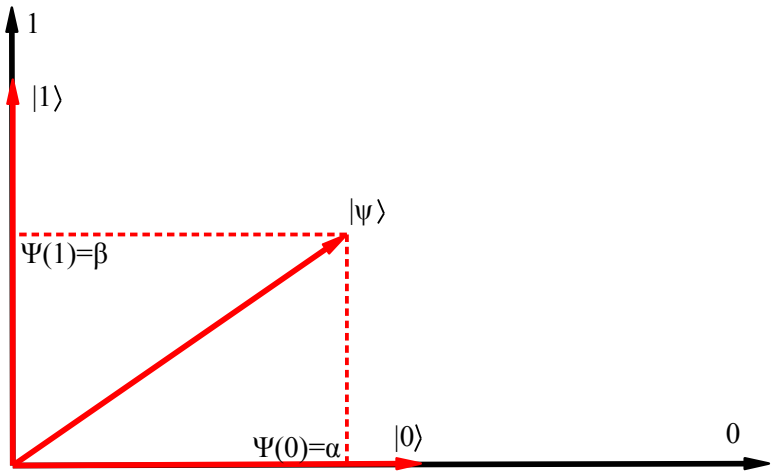
So, a ket vector is denoted by the symbol $|\rangle$ and the title within this symbol. Each state of a dynamical sytem is associated to a different ket vector. As we know, a linear superposition of states of a quantum system is also a possible state of the system.

From this fact, it follows that the space of ket vectors is a linear one, in the following sense: if c_1 and c_2 are complex numbers and $|a\rangle$ and $|b\rangle$ are ket vectors, their linear combination is also a ket vector:

$$|u\rangle = c_1|a\rangle + c_2|b\rangle,$$

because the linear combination of states $|a\rangle$ and $|b\rangle$ is also a state of the system. If we have several ket vectors and none of them can be expressed as a linear combination of other ones, our vectors are said to be linearly independent vectors. Number of dimension of a ket vector space is defined by the number of linearly indenpendent vectors contained by the vector space.

Independent states (which are mutually excluded by each other, in classical physics) of a quantum system are represented by linearly independent ket vectors, hence number of dimension of (state-) space of these ket vectors is defined by the number of indenpendent states of the quantum system. In this interpretation (or picture), time evolution of the system can be imagined as a rotation of the state vector around the origo. (Let us try to remember, this is exactly the Schrödinger picture discussed in the introductory chapter.) Returning to qubits: we remember that any two-state quantum system can be considered as a qubit, where the state of the system is the information itself. To make its visualization easier, we try to illustrate a quantum bit (or rather its state vector) in the picture below:



In the picture above, qubit state $|\Psi\rangle$ is a superposition of a state titled by 0 as $|0\rangle$ and a state titled by 1 as $|1\rangle$, where states $|0\rangle$ and $|1\rangle$ take part with weights α and β. It is important for us to know α and β can be complex numbers. (The ones who

followed materials from the beginning of this semester can surmise how to obtain probabilities of choosing state $|0\rangle$ or state $|1\rangle$ by the qubit, if a measurement is made.) Being a two-state system, in case of observation (measurement) a qubit can choose from between maximum two, independent states. Polarization of a photon, or spin of a particle, etc belong to systems of this kind. For example, in case of spins, state "spin up" ($|0\rangle$) can be associated to bit 0, and state "spin down" ($|1\rangle$) can be associated to bit 1. Using the base expanded by kets $|0\rangle$ and $|1\rangle$, it is important for us to know how to calculate anything we want to know about the system. Being afraid of forgetting to mention, we have to mention a very important knowledge, namely: *length of a ket vector has to be* 1*, always*[1]. Obviously in this base, on the axe associated to bit 0, value of coordinate of ket $|0\rangle$ equals 1, and on the other axe its coordinate equals 0 (and vice versa in case of ket $|1\rangle$). Ket vectors are generally represented by column matrix. In this case they can be written into the following shape:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We will need to use also adjoints of ket vectors (adjoint means: transposed and (complex) conjugated). These kind of vectors are named *bra* vectors. In case of having non-complex components, adjoints of kets $|0\rangle$ and $|1\rangle$ (bras $\langle 0|$ and $\langle 1|$) can be written into the form below:

$$\langle 0| = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \langle 1| = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

Scalar (or in other word: inner) product of two vectors will also be needed. This multiplication can be achieved in the following way: we have the adjoint of one of the two vectors (in case of non-complex representatives, it is all the same which vector is adjoined), then we multiply this adjoint vector by the other vector. For example scalar product of kets $|0\rangle$ and $|1\rangle$ can be seen below:

$$\langle 1|0\rangle = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

In case of non-complex components, order of factors does not matter, because:

[1] more on this later

$$\langle 1|0\rangle = \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \langle 0|1\rangle = \begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = 0,$$

but on the other hand, in case of kets $|a\rangle$ and $|b\rangle$ have complex representatives, the following formula is valid:

$$\langle a|b\rangle = \langle b|a\rangle^*$$

....and of course, we will need to use also the outer product of vectors. Instead of a number, this kind of product results a matrix. As we can see below, in case of outer product, order of factors is not irrelevant, even in the case of non-complex components:

$$|0\rangle\langle 1| = \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} = \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}$$

$$|1\rangle\langle 0| = \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}$$

Now, let us return to our quantum bit:

$$|0\rangle, \quad |1\rangle \Longrightarrow |\Psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

where the following very important requirement has to be satisfied:
$|\alpha|^2 + |\beta|^2 = 1$. (This requirement is necessary because of the probability interpretation.) In this phase, we have to enlighten the reasons why we have to know the things treated above. First of all, as we saw, state $|\Psi\rangle$ of a quantum bit is expressed on an orthonormal base which is expanded by kets $|0\rangle$ and $|1\rangle$. States that are represented by kets $|0\rangle$ and $|1\rangle$ are called classically observable states. These orthonormal, *observable*[2] states are the *eigenstates* of the system. If a measurement is made, the measured system falls always into one of its eigenstates (in case of our qubit, into ket $|0\rangle$ or into ket $|1\rangle$) and from that instant, state vector of the system will be the "chosen" eigenstate (or eigenvector). Now then! The question is that how can we calculate the probability of finding the system in a certain eigenstate, after making a measurment? We can calculate it via ascertaining the

[2]or in other words: states that are mutually excluded by each other in classical physics

amplitude of projection of state vector $|\Psi\rangle$ on the certain eigenstate[3]. After obtaining this projection (which is a complex number), we have to compute its absolute value, and then we have to square this absolute value. The calculated number will be the value of the sought probability.

For example, if we want to know the probability of finding the system of qubit $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$ in the state $|0\rangle$, after a measurement, we have to calculate the square of absolute value of α. Square of absolute value of a complex number equals the number itself multyplied by its complex conjugated. In our case, it means: $|\alpha|^2 = \alpha^*\alpha$. Since we will work with vectors with two or far more then two number of dimensions, it is necessary to know how to calculate probabilities that we are interested in. First of all, let us keep in mind that scalar product of two vectors which are perpendicular to each other equals 0, furthermore let us realize that all eigenvectors in the linear combination which builds up state $|\Psi\rangle$ are mutually perpendicular to each other (in our case: $|0\rangle$ and $|1\rangle$). From these two facts it directly follows the way of calculation of probabilities which we are interested in:

We have to square the absolute value of scalar product of state $|\Psi\rangle$ and the eigenstate whose materialization-probability we want to know. This method can be understood very easy, if we keep in mind that $|\Psi\rangle$ is a sum of mutually *perpendicular* (eigen)vectors, thus in case of multiplying it by *a member of this sum*, we actually do the following: we multiply all members of the sum by the given member and add the obtained results to each other. Consisting of mutually perpendicular ket vectors, all scalar products which contains different kets results 0. Only the scalar product survives this operation in which the angle between the eigenvector and our ket is equal to 0. There is only one such an eigenvector in the linear combination of $|\Psi\rangle$, namely the one by which $|\Psi\rangle$ was multiplied. Obviously, scalar product between our eigenket and itself equals 1, and this result multiplies the *projection* which we are intersted in (which can be called *probability amplitude*).

What is it all about? If we are interested in the value of the projection of $|\Psi\rangle$ on a certain eigenket, all we have to do is to have the scalar product between $|\Psi\rangle$ and the eigenket. Let us consider a very simple example, where we show how to compute the value of probability w_1 of finding the system in state $|1\rangle$ after making a measurement on the base of ($|0\rangle$, $|1\rangle$), if the initial state of the system is $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$:

$$w_1 = |\langle 1|\Psi\rangle|^2 = |\langle 1|(\alpha|0\rangle + \beta|1\rangle)|^2 = |\alpha\langle 1|0\rangle + \beta\langle 1|1\rangle|^2 =$$

$$= |\alpha \times 0 + \beta \times 1|^2 = |\beta|^2 = \beta^*\beta$$

[3] obviously *before* making the measurement

...or in matrix representation:

$$w_1 = |\langle 1|\Psi\rangle|^2 = |\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}|^2 = |\beta|^2 = \beta^*\beta$$

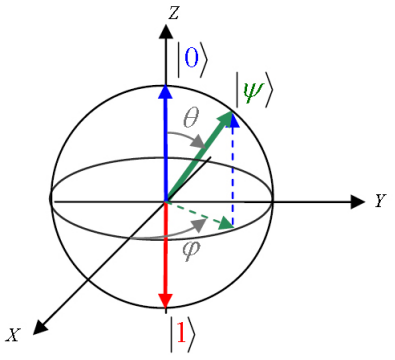
Considering things drawn above, we can understand the reason of the requirement connected with $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$, namely: $|\alpha|^2 + |\beta|^2 = 1$. Since $|\alpha|^2$ is the probability of finding the system in state $|0\rangle$, and $|\beta|^2$ is the probability of materialization of state $|1\rangle$ after a measurement, obviously their sum has to be equal to 1. From this fact, it follows that *the length (or norm) of any statevector has to equal* 1, (because in case of $|\alpha|^2 + |\beta|^2 = 1$, length $\sqrt{|\alpha|^2 + |\beta|^2}$ will be equal to 1). From the foregoing, it is clear how to calculate the norm (length) of an arbitrary vector $|\Psi\rangle$. (By the way, the norm of a vector $|\Psi\rangle$ is denoted by $\|\Psi\|$.)

$$\begin{aligned}\|\Psi\| = \sqrt{\langle\Psi|\Psi\rangle} = \sqrt{(\alpha^*\langle 0| + \beta^*\langle 1|)(\alpha|0\rangle + \beta|1\rangle)} = \\ = \sqrt{|\alpha|^2\langle 0|0\rangle + \alpha^*\beta\langle 0|1\rangle + \beta^*\alpha\langle 1|0\rangle + |\beta|^2\langle 1|1\rangle} = \\ = \sqrt{|\alpha|^2 + |\beta|^2}\end{aligned}$$

...or it can be expressed via matrices:

$$\|\Psi\| = \sqrt{\langle\Psi|\Psi\rangle} = \sqrt{\begin{pmatrix}\alpha^* & \beta^*\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}} = \sqrt{|\alpha|^2 + |\beta|^2}$$

Generally, α is a real number and β is a complex one. Besides superposition, phase-freedom is also a newness, in quantum information (from the $e^{i\varphi}$ shape of complex numbers). Phase-freedom makes possible for us to achieve computations via interference.



This is related to another kind of visualization of qubits which is represented by the *Bloch-sphere*, as it can be seen in the picture right. In this case, the state of the qubits is described by two angles ($0 \leq \theta \leq \pi$ and $0 \leq \varphi \leq 2\pi$), according to the following formula: $|\Psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle$. Any point on the surface of the sphere is a possible state $|\Psi\rangle$. In this kind of representation θ is responsible for superposition. If $\theta = 0$, the system is in state $|0\rangle$, and in case of $\theta = \pi$, state of the system is $|1\rangle$. On this surface, orthogonal

states are 'over against' each other. Let us remember, in our previous representation, orthogonal states were perpendicular to each other. This is the reason why $\frac{\theta}{2}$ is written into the expression above, instead of θ. Angle φ represents the phase of the complex factor.

4.1.2 Description of a multi-qubit system

Naturally, we will have to work with systems which consist of several qubits, thus we need to know the way of description of their state. Firstly we stay at the case of ensemble of two qubits. In this case the - four dimensional - state space of their common system is the tensor product of their - two dimensional - state spaces. The base where their state is expressed on is the so called product base, whose elements are the following ones:

$$|0\rangle_1|0\rangle_2 = |00\rangle$$
$$|0\rangle_1|1\rangle_2 = |01\rangle$$
$$|1\rangle_1|0\rangle_2 = |10\rangle$$
$$|1\rangle_1|1\rangle_2 = |11\rangle$$

In our case $|\Psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$, and obviously the following requirement has to be satisfied: $|\alpha_{00}|^2 + |\alpha_{01}|^2 + |\alpha_{10}|^2 + |\alpha_{11}|^2 = 1$. Generally one of the α-s gets a real value while all the others are complex, hence we get a very large phase-freedom. Let us realize that numbers in the symbols of kets above are actually binary numbers. In decimal system, they correspond to numbers 0, 1, 2, and 3. In case of a system which consists of three qubits, the first base ket of the state space is $|000\rangle$ and the last one is $|111\rangle$, thus numbers from 0 to 7 can be described by base kets. If we have an *n*-qubit system, base elements of its state space are the following ones:

$$|00.....00\rangle$$
$$.$$
$$.$$
$$.$$
$$|11.....11\rangle$$

From this, it follows that general state of an *n*-qubit system can be written into the following shape:

$$|\Psi\rangle = \sum_{x=0}^{2^n-1} c_x |x\rangle$$

The question naturally arises how can a tensor production of vectors be achieved?[4] Without any explanation we try to show this method below via an example:

$$|a\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad |b\rangle = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad |a\rangle|b\rangle = |ab\rangle = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ a_2 b_1 \\ a_2 b_2 \end{pmatrix}$$

In case we have non column/row matrices:

$$a = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}, \quad a \otimes b = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_2 b_1 & a_2 b_2 \\ a_1 b_3 & a_1 b_4 & a_2 b_3 & a_2 b_4 \\ a_3 b_1 & a_3 b_2 & a_4 b_1 & a_4 b_2 \\ a_3 b_3 & a_3 b_4 & a_4 b_3 & a_4 b_4 \end{pmatrix}$$

Returning to *n*-qubit systems, we have to mention possibilities given by them: as we have already known numbers and *their superpositions* can be described by two-state quantum systems. What is it all about? As we know $|\Psi\rangle = \sum_{x=0}^{2^n-1} c_x |x\rangle$ is the general state of an *n*-qubit system, where in each $|x\rangle$, x is binary number consisting of n pieces of digits. In a *superposition* of an *n*-qubit system, **many numbers** ($2^n - 1$ and 0) can be described **at the same time** via *one* of the states of the system (and of course there is the phase-freedom, as an additional possibility).

[4] For example: how to achieve a production like this $|00\rangle$?

Chapter 5

Quantum gates

In the first part of this semester, from *Schrödinger*-equation we derived an important conservation law related to the current of probability-density. The derived law of conservation represents the conservation of the total probability. It means that any transformation which has an effect on our state vectors (for example in case of their time evolution) must leave the total probability untouched. What does it mean exactly? The thing is that though state vector may rotate around the origo as the state itself is changing, its length has to be left unchanged by the transformation which makes state vector rotate.

Why is it equivalent to the conservation of probability? Let us consider the well known two dimensional case, where the norm (length) of the state vector $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$ was calculated according to the formula below:

$$\|\Psi\| = \sqrt{\langle\Psi|\Psi\rangle} = \sqrt{|\alpha|^2 + |\beta|^2}$$

Members under the root-sign are values of probabilities. Any change of the length of state vector under some transformation would mean that sum of probabilities has changed, or in other words: total probability is not invariant. But - as we know - it cannot happen because of the law of probability-conservation. Since **quantum gates** alter states, matrices which describe transformations made by these gates has to represent operators which satisfy the requirement mentioned above. Operators of this kind are called **unitary operators** (and their matrices are unitary ones). Unitarity is a necessary and sufficient condition which ensures that transformation cannot change the length of (the norm of) a state vector. Operators of such kind of transformations are denoted by the following symbol: $\hat{U}$. But

what is the condition of unitarity? A matrix which represents a unitary operator has to satisfy the following requirement: $U^{-1} = U^{\dagger}$, where $U^{\dagger}$ means the adjoint (transposed and complex conjugated) of the matrix.

From the fact that length of our state vectors $\sqrt{\langle\Psi|\Psi\rangle} = 1$ (because of the total probability), it follows that $\langle\Psi|\Psi\rangle = 1$. If $|\Psi\rangle$ is subject to a unitary operation of some $\hat{U}$ which alters the original state into a new state of $|\Psi'\rangle$, the following expression has to come true: $\langle\Psi|\Psi\rangle = \langle\Psi'|\Psi'\rangle = 1$. Let us check how unitarity of an operator $\hat{U}$ can ensure it:

$$|\Psi'\rangle = \hat{U}|\Psi\rangle$$
$$\langle\Psi'|\Psi'\rangle = \langle\hat{U}\Psi|\hat{U}\Psi\rangle = \langle\Psi|\hat{U}^{\dagger}\hat{U}|\Psi\rangle$$

We know that according to unitarity $U^{-1} = U^{\dagger}$, furthermore we know U is a square matrix, and in case of a square matrix $UU^{-1} = U^{-1}U = I$. Knowing these facts, we can realize the value of the scalar product (just like the value of the norm) remained the same after the effect of $\hat{U}$. There is another important consequence of unitarity of an operator like $\hat{U}$, namely: transformations described by this kind of operators are *reversible*. This means output state of a quantum gate (or a quantum circuit containing several gates) can be retransformed to its input state by sending back the output state into the gate (or circuit).

Let us consider what it means in case of *one* quantum gate:
So, our input sate is $|\Psi\rangle$, and state $|\Psi'\rangle = U|\Psi\rangle$ will appear on the output. (Since henceforward every operator will be represented by matrices, from now $\hat{U} = U$.) This output state will be sent back through the quantum gate, in other words matrix U will have an effect on state $|\Psi'\rangle = U|\Psi\rangle$. In this way, we will get the following state: $UU|\Psi\rangle$.

At this point, it is important to know that from the unitarity of a matrix U, it follows U is a self-adjoint (hermitian) matrix, thus $U^{\dagger} = U$. Accordingly, $UU|\Psi\rangle = U^{\dagger}U|\Psi\rangle$. Since - as we know - $U^{\dagger} = U^{-1}$, our last expression can be altered into the next form: $U^{\dagger}U|\Psi\rangle = U^{-1}U|\Psi\rangle = |\Psi\rangle$. As we can see, we got back the original input state, thus the transformation is really reversible. From this reversibility, it follows there are classical gates which do not have a quantum analogous. For example such a gate is the AND gate which gives 0 output in case of 00, 01, and 10 inputs. In such a case, input cannot be found out from knowledge of output. On the other hand, NOT gate, which causes a bit-flip ($0 \rightarrow 1$, $1 \rightarrow 0$), is a reversible one, thus it has a quantum analogous ($|0\rangle \rightarrow |1\rangle$, $|1\rangle \rightarrow |0\rangle$). Its effect on a general qubit is:

$$\alpha|0\rangle + \beta|1\rangle \rightarrow \alpha|1\rangle + \beta|0\rangle$$

Representing the state of the qubit by a two-component column matrix like this

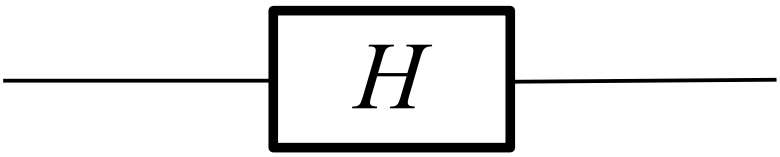
$$\alpha|0\rangle + \beta|1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

we can describe the effect of NOT gate by use of the matrix below:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix},$$

which is actually one of the three *Pauli* matrices[1] denoted by σ_x.

Now, that we know there are classical gates which do not have quantum analogous (for example: AND gate), it is time to mention a quantum gate that does not have classical analogous. Such a gate is the one-bit *Hadamard* gate whose circuit symbol can be seen below:



The left side line of the picture above represents the input quantum bit, and the right side symbolizes the output qubit. It is nice, but we still have not written what happen with qubit crossing through the *Hadamard* gate. Depending on the bit itself, the gate can have two kinds of effect on the qubit:

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

(Let us realized that the resulted states span an orthonormal base.) So, after the effect of *Hadamard* gate, a qubit which was originally in one of its eigenstates

[1] more on them later

(in $|0\rangle$ or $|1\rangle$) gets into the superposition of states $|0\rangle$ and $|1\rangle$. It is apparently *impossible in a classical case*. Let us realize that $H^2 = I$:

$$HH|0\rangle = \frac{1}{\sqrt{2}}H(|0\rangle + |1\rangle) = \frac{1}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) + \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right] =$$

$$= \frac{1}{\sqrt{2}}\frac{2}{\sqrt{2}}|0\rangle = |0\rangle$$

The matrix which represents the *Hadamard* gate can be seen below:

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Naturally, there are two-qubit gates, too. From this kind of gates, C-NOT (Controlled–NOT) gate is the most important one, whose effect can be found below:

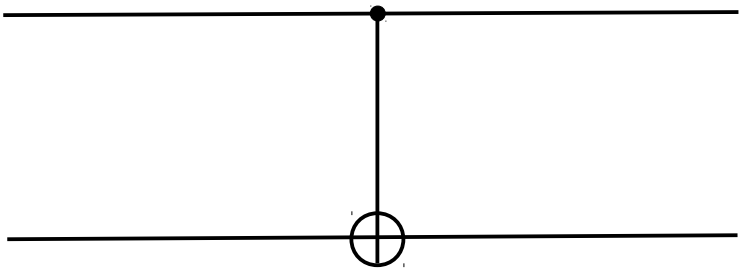
$$|00\rangle \rightarrow |00\rangle$$
$$|01\rangle \rightarrow |01\rangle$$
$$|10\rangle \rightarrow |11\rangle$$
$$|11\rangle \rightarrow |10\rangle$$

From the rows above, it is clear that C-NOT affects only in case value of first qubit (the control bit) equals 1. In this case second qubit (target bit) is given a reverse value than it had before (a bit flip happens). If we have a little patient, we can easily find out the matrix which represents the effect of C-NOT:

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Watchful readers can find quickly the NOT gate hiding in the right bottom part of the matrix. Checking of unitarity of C-NOT is a reader's task.

Circuit symbol of C-NOT can be seen below:



In the picture above, upper line symbolyzes the controll bit whose value remains untouched by the gate. Only the value of target bit (visualized by the lower line) can change, in case value of control bit equals 1.

5.1 Quantum circuits

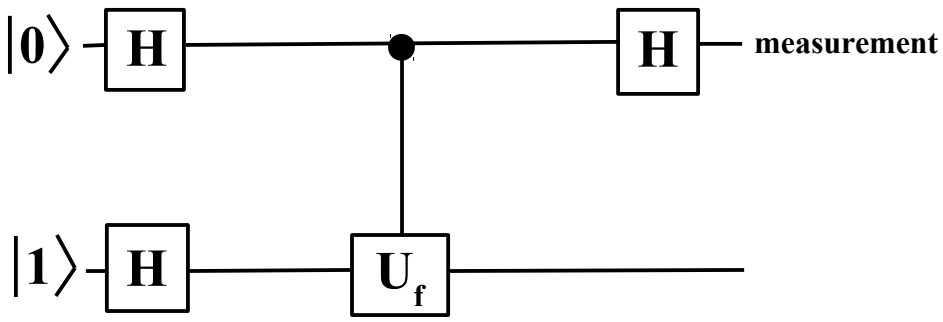
Now we have reached the point, where it is worth introducing the method of modelling quantum computation via quantum circuits. As it can be surmised, in a system of this kind of circuits, qubits are symbolized by lines of the circuit, and quantum gates (which are actually unitary opertors) are denoted by their symbols. There is an important theorem, namely: Any kind of unitary transformation can be constructed by use of one-qubit gates and C-NOT gates. (We set aside from proving this theorem.) This is the reason why we can meet so many C-NOT gates in schemes realizing different kinds of quantum informatical protocols. Below, we show a very simple quantum protocol.

5.1.1 Deutsch algorithm

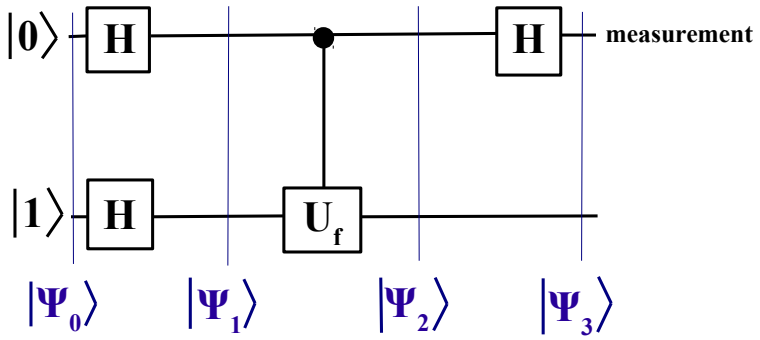
As a tradition, many times *Deutsch* problem is the first treated one in standard textbooks on quantum information to demonstrate how efficient a quantum algorithm can be. The reason which makes writers of these books follow this way can be easily understood, if we realize that though using very few knowledges, this algorithm already shows quantum paralellism in computation, giving students a relatively simple tool that demonstrates how a quantum algorithm can work tipically. Let us follow our great ancestors by beginning with this algorithm.

Let us consider a function f which maps the discreet set of $\{0, 1\}$ onto the set $\{0, 1\}$. In case $f(0) = f(1)$, function f is called constant, and if $f(0) \neq f(1)$

function f is a balanced one. Using some classical method, we have to make *two* evaluations to decide whether function f is a constant or a balanced one. However, using Deutsch algorithm, we can decide it after *one* evaluation. Let us check how it works. First of all, we have to know the quantum circuit which implements Deutsch's algorithm. This circuit is drawn in the figure below:



In the picture above, U_f denotes the f-Controlled-NOT gate (or fNOT, CfNOT, or Deutsch gate). What does this gate do? In case of denoting the upper bit (which is the control bit) by $|x\rangle$, and the lower bit (target bit) by $|y\rangle$, two-qubit gate U_f has the following effect on the pair of qubits: $|x\rangle|y\rangle \rightarrow |x\rangle|y \oplus f(x)\rangle$, where $\oplus$ means modulo-2 addition, which is subject to the following connections: $0 + 0 = 0$, $0 + 1 = 1$, $1 + 0 = 1$, $1 + 1 = 0$. In the figure below, different states of the two-qubit system is denoted, correspondently to relevant sections of the circuit.



As it can be seen, input state of the circuit is $|\Psi_0\rangle = |0\rangle|1\rangle$. Due to the effect of first two *Hadamard* gates, this state alters into the state $|\Psi_1\rangle$, as we can see below:

$$|\Psi_1\rangle = H|0\rangle H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) =$$

$$= \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle)$$

Now let us see how does this state change due to effect of Deutsch gate:

$$|\Psi_2\rangle = CfNOT|\Psi_1\rangle = \frac{1}{2}CfNOT(|00\rangle - |01\rangle + |10\rangle - |11\rangle) =$$

$$= \frac{1}{2}[|0\rangle|0 + f(0)\rangle - |0\rangle|1 + f(0)\rangle + |1\rangle|0 + f(1)\rangle - |1\rangle|1 + f(1)\rangle] =$$

$$= \frac{1}{2}[|0\rangle(|0 + f(0)\rangle - |1 + f(0)\rangle) + |1\rangle(|0 + f(1)\rangle - |1 + f(1)\rangle)]$$

In case of being watchful, we can realize some possibilities which can simplify the expression above, namely:

- if $f(x) = 0$, then $|0 + f(x)\rangle - |1 + f(x)\rangle = |0\rangle - |1\rangle$
- on the other hand,
  if $f(x) = 1$, then $|0 + f(x)\rangle - |1 + f(x)\rangle = |1\rangle - |0\rangle = (-1)(|0\rangle - |1\rangle)$
- *thus*: $|0 + f(x)\rangle - |1 + f(x)\rangle = (-1)^{f(x)}(|0\rangle - |1\rangle)$

From this enumeration, it follows that $|\Psi_2\rangle$ can be written into the shape below:

$$|\Psi_2\rangle = \frac{1}{2}[(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle](|0\rangle - |1\rangle)$$

Before making the measurement (in other word: readout) on the output of the circuit, we have to find out the effect of the last gate on the state of the system. As we can see, in this case a *Hadamard* gate affects on the upper qubit. Let us calculate the resulted state after its action:

$$|\Psi_3\rangle = \frac{1}{2\sqrt{2}}[(-1)^{f(0)}(|0\rangle + |1\rangle) + (-1)^{f(1)}(|0\rangle - |1\rangle)](|0\rangle - |1\rangle) =$$

$$= \frac{1}{2\sqrt{2}}[((-1)^{f(0)} + (-1)^{f(1)})|0\rangle + ((-1)^{f(0)} - (-1)^{f(1)})|1\rangle](|0\rangle - |1\rangle)$$

So, $|\Psi_3\rangle$ is the output state of the circuit, and as it can be seen, there remains no more but to make a measurement on this output state to find out whether function f is a constant or a balanced one. It is natural to ask: how can we decide it based on only one measurement (or in other word: evaluation)? The answer can be easily understood, if we try to imagine what is the output state $|\Psi_3\rangle$ in case of a constant function and in case of a balanced function. Let us see the result of our imagination:

- if f is a constant function, the resulted state is

$$|\Psi_3\rangle = \pm\frac{1}{\sqrt{2}}(|00\rangle - |01\rangle)$$

- if f is a balanced function, the resulted state is

$$|\Psi_3\rangle = \pm\frac{1}{\sqrt{2}}(|10\rangle - |11\rangle)$$

So, we can draw a conclusion, according to which if f is a constant function, state of the upper qubit will be $|0\rangle$, and if f is a balanced one, upper qubit will be in the state $|1\rangle$.

Thus, making *one* measurement only on the upper qubit, we can decide what sort of function f is.

Chapter 6

Density matrix, mixed state

Density matrix has already been partly explained in the itroductory chapter. Now we detail some of its important features which will be useful later. First of all we have to define a very considerable idea, namely: *ensemble* which is a set of similar quantum systems prepared in different states. Suppose we select one of these systems. The probability of selecting a system whose state is Ψ_i equals p_i and $\sum_{i=1}^{n} p_i = 1$, where n denotes the number of the systems. Measuring a physical quantity A, we would like to know its expectation value on this ensemble. In other words, we want to know $\langle \hat{A} \rangle$. In a state $|\Psi_i\rangle$ – as we know – the expectation value of A can be calculated in the following way: $\langle \Psi_i | \hat{A} | \Psi_i \rangle$. However we need to average twice: besides the former expression, we have to average over weights too. Hence

$$\langle \hat{A} \rangle = \sum_{i=1}^{n} p_i \langle \Psi_i | \hat{A} | \Psi_i \rangle = Tr(\hat{\varrho} \hat{A}),$$

where $\hat{\varrho}$ is the well known density operator which features the ensemble and as we can surmise $\hat{\varrho} = \sum_{i=1}^{n} p_i |\Psi_i\rangle\langle\Psi_i|$. How did we get this expression of $\hat{\varrho}$? Let us consider the followings:

As we know:

$$|\Psi_i\rangle = \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha_i \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \beta_i \end{pmatrix} = \alpha_i \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta_i \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \alpha_i |1\rangle + \beta_i |1\rangle.$$

Let us take the inner product between $|\Psi_i\rangle$ and $|\Psi_j\rangle$:

$$\langle\Psi_i|\Psi_j\rangle = \begin{pmatrix}\alpha_i^* & \beta_i^*\end{pmatrix}\begin{pmatrix}\alpha_j \\ \beta_j\end{pmatrix} = \alpha_i^*\alpha_j + \beta_i^*\beta_j$$

Let us mark this result and consider the outer product of these kets:

$$|\Psi_i\rangle\langle\Psi_j| = \begin{pmatrix}\alpha_i \\ \beta_i\end{pmatrix}\begin{pmatrix}\alpha_j^* & \beta_j^*\end{pmatrix} = \begin{pmatrix}\alpha_i\alpha_j^* & \alpha_i\beta_j^* \\ \beta_i\alpha_j^* & \beta_i\beta_j^*\end{pmatrix}$$

As we can see, having the trace of the outer product (summing the diagonal element of the matrix above), we obtain exactly the result of the inner product[1].

So

$$Tr(|\Psi_i\rangle\langle\Psi_j|) = \sum_n \langle n|\Psi_i\rangle\langle\Psi_j|n\rangle = \sum_n \langle\Psi_j|n\rangle\langle n|\Psi_i\rangle = \langle\Psi_j|\hat{I}|\Psi_i\rangle = \langle\Psi_j|\Psi_i\rangle.$$

That is the operation of the trace makes the outer product an inner product. Returning to the initial statement:

$$Tr\left[\hat{A}\sum_i p_i|\Psi_i\rangle\langle\Psi_i|\right] = \sum_i p_i Tr\left[\hat{A}|\Psi_i\rangle\langle\Psi_i|\right] = \sum_i p_i\langle\Psi_i|\hat{A}|\Psi_i\rangle$$

So the density operator describes an ensemble. Nevertheless, there is another interpretation of the density operator, namely: when a subsystem of a larger system is considered. As an example, let us suppose we have two systems, A and B. In this case the common state vector is $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, which can be expressed on the following base: $\big\{\{|n\rangle_A\};\quad \{|m\rangle_B\}\big\}$. Hence $|\Psi\rangle = \sum_{nm} c_{nm}|n\rangle|m\rangle$ and $\sum_{nm}|c_{nm}|^2 = 1$. Let us consider a system X_A, which is one of the physical quantities of system A. In this situation, $\hat{X}_A \otimes \hat{I}_B$ is a two-system operator measuring X_A in system A, but leaving B untouched. Its expectation value is

$$\langle\hat{X}_A\rangle = \langle\Psi|\hat{X}_A\otimes\hat{I}_B|\Psi\rangle = \sum_n\sum_m \langle\Psi|\hat{X}_A\otimes\hat{I}_B\big(|n\rangle_A|m\rangle_{B\ A}\langle n|\ {}_B\langle m|\big)|\Psi\rangle = \sum_n {}_A\langle n|\sum_m {}_B\langle m|\Psi\rangle\langle\Psi|m\rangle_B\hat{X}_A|n\rangle_A,$$

[1]The only difference is that i is replaced with j.

where

$$\sum_m {}_B\langle m|\Psi\rangle\langle\Psi|m\rangle_B$$

is the reduced density operator of subsystem A denoted by $\hat{\varrho}_A$. In the derivation $\sum_n {}_A\langle n|\sum_m {}_B\langle m|\Psi\rangle$ is just an expression factor and $\langle\Psi|m\rangle_B$ was also replacable, because $\hat{X}_A$ does not have an effect on $|m\rangle_B$, because it is in system B.)

Knowing that

$$\hat{\varrho}_A = Tr_B|\Psi\rangle\langle\Psi| = \left(\sum_m {}_B\langle m|\Psi\rangle\langle\Psi|m\rangle_B\right),$$

we have that

$$\langle\hat{X}_A\rangle = Tr_A(\hat{\varrho}_A\hat{X}_A).$$

In this case, though the larger system is in a pure state, its subsystem is in a mixed state.

Let us consider examples for both interpretations:

Example for the ensemble interpretation (in case of qubits):

Let us suppose we many qubits and half of the qubits are in state $|0\rangle$, while the other half of them are in state $|1\rangle$. In this case the density operator is

$$\hat{\varrho} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\hat{I}.$$

Let us try to find out the expectation value of σ_z. As we know it is one of the three Pauli operator represented by the following matrix:

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

By the way $|0\rangle$ and $|1\rangle$ are the eigenstates of σ_z with eigenvalues 1 and -1:

$$\sigma_z|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle$$

$$\sigma_z|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -|1\rangle$$

From this it follows that

$$\langle \sigma_z \rangle = \frac{1}{2}1 + \frac{1}{2}(-1), \text{ which is obviously equals } Tr(\hat{\varrho}\sigma_z) = 0.$$

Example for a system consisting of two qubits in a state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B).$$

We do not know it *yet*[2] but this is one of the four famous *Bell* states:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle + |1\rangle|0\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

Now let us see the spin of A. We know that $Tr_B(|\Psi\rangle\langle\Psi|) = \hat{\varrho}_A$, thus

$$\hat{\varrho}_A =$$
$$\sum_{m=0}^{1} {}_B\langle m| \Big\{ \frac{1}{\sqrt{2}} \big(|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B \big) +$$
$$\frac{1}{\sqrt{2}} \big({}_A\langle 0|\, {}_B\langle 1| + {}_A\langle 1|\, {}_B\langle 0| \big) \Big\} |m\rangle_B =$$
$$\frac{1}{2} \sum_{m=0}^{1} {}_B\langle m| \Big\{ |0\rangle_A|1\rangle_B\, {}_A\langle 0|\, {}_B\langle 1| + |0\rangle_A|1\rangle_B\, {}_A\langle 1|\, {}_B\langle 0| +$$
$$|1\rangle_A|0\rangle_B\, {}_A\langle 0|\, {}_B\langle 1| + |1\rangle_A|0\rangle_B\, {}_A\langle 1|\, {}_B\langle 0| \Big\} |m\rangle_B =$$

[2] But it will be shortly discussed.

$$\frac{1}{2}\big(|0\rangle_{A\,A}\langle 0| + |1\rangle_{A\,A}\langle 1|\big).$$

Since $\sigma_{z_A} = \sigma_z \otimes \hat{I}_B$,

$$\langle \sigma_{z_A} \rangle = Tr(\sigma_z \varrho_A) = 0.$$

In this formalism, if we want, also a pure state can be described by a density operator, where

$$\varrho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| = |\Psi\rangle\langle\Psi|$$

because each element of the ensemble is in the same state, whose weight is 1. Weight belonging to all the other states is 0. The expectation value of a physical quantity A is

$$\langle \hat{A} \rangle = Tr(\hat{\varrho}\hat{A}) = Tr(|\Psi\rangle\langle\Psi|\hat{A}) = \langle\Psi|\hat{A}|\Psi\rangle.$$

Both pure and mixed states can be featured by density operators. If a density operator has a shape like this $\varrho = |\Psi\rangle\langle\Psi|$, the state is said to be a pure state, but in case of weighted sum of pure states – $\varrho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|$ – the related state is a mixed state. It is natural to ask that in case we have a density operator, how can we find out if it is a pure or a mixed state? This question can be answered with the properties of density operators:

- $$Tr(\hat{\varrho}) = 1$$

  Proving:

  $$Tr(\hat{\varrho}) = Tr\left(\sum_i p_i |\Psi_i\rangle\langle\Psi_i|\right) = \sum_i p_i Tr\left(|\Psi_i\rangle\langle\Psi_i|\right) =$$

  $$\sum_i p_i \langle\Psi_i|\Psi_i\rangle = \sum_i p_i = 1$$

- $$\hat{\varrho} = \hat{\varrho}^\dagger$$

This feature comes from the construction of the density operator, because $|\Psi_i\rangle$ and $\langle\Psi_i|$ are the adjoints of each other. From this feature, it follows that eigenvalues of $\hat{\varrho}$ are real.

- $\hat{\varrho}$ is a positive operator, that is $\langle\Psi|\hat{\varrho}|\Psi\rangle \geq 0$. Proving:

$$\langle\Psi|\hat{\varrho}|\Psi\rangle = \langle\Psi|\left(\sum_i p_i|\Psi_i\rangle\langle\Psi_i|\right)|\Psi\rangle = \sum_i p_i\langle\Psi|\Psi_i\rangle\langle\Psi_i|\Psi\rangle =$$

$$\sum_i p_i|\langle\Psi_i|\Psi\rangle|^2 \geq 0$$

- $\hat{\varrho}$ is positive if and only if all of its eigenvalues are $\lambda_i \geq 0$. Being a Hermitian operator, it has a diagonal production: $\hat{\varrho} = \sum_i \lambda_i|u_i\rangle\langle u_i|$, where it is true that $\langle u_i|u_j\rangle = \delta_{ij}$ and $\sum_i |u_i\rangle\langle u_i| = \hat{I}$, ($\lambda_i \geq 0$). Though different $|\Psi_i\rangle$-s are not necessarily perpendicular to each other diagonalization can be made (spectral theorem).

  $\hat{\varrho}$ is a pure state if and only if

$$tr(\hat{\varrho}^2) = 1.$$

  Obviously, in case $\hat{\varrho}$ is a pure state:

$$\hat{\varrho} = |\Psi\rangle\langle\Psi| \Rightarrow \hat{\varrho}^2 = \big(|\Psi\rangle\langle\Psi|\big)\big(|\Psi\rangle\langle\Psi|\big) = |\Psi\rangle\langle\Psi| = \hat{\varrho}$$

  (let us remember: $\langle\Psi|\Psi\rangle = 1$), and the trace of $\hat{\varrho}$ is 1.

Without additional provings, we declare that in case of a mixed state $Tr(\hat{\varrho}^2) < 1 = Tr(\hat{\varrho})$ and for a pure state $Tr(\hat{\varrho}^2) = 1 = Tr(\hat{\varrho})$.

Chapter 7

A quantum cryptographical protocol

So far, we have learned the simplest tools which are necessary for us to understand the most elementary algorithms and methods in quantum information. Hence, by this time, we have become clever enough :-) to begin to deal with some of the simplest applications of the easiest parts of quantum mechanical toolkit in the area of quantum communication. Nevertheless, some algorithms and methods belonging to the same area have to be treated separately, because even now, we do not have knowledge enough in quantum mechanics. What is it all about? For understanding this dilemma, let us consider - for example - the quantum cryptographical protocols. These methods should be explained in the same chapter, accordingly their goals. But we can not do this, for several reasons. Just as an example of these reasons: there is a cryptographical protocol called E91 which is based on the phenomenon of quantum entanglement, and if we would like to understand how it works, we have to know the fundamental theoretical background of the entanglement. Being an important topic which is unknown for us, entanglement requires an own chapter, thus it should be treated before learning about certain cryptographical – and some other – protocols.

Nevertheless we would like to maintain the attendance of students, hence we start to learn about *practical applications* which can be already understood using the discussed / known theoretical tools, and those applications which still can not be understood will be discussed in a future chapter after explaining their theoretical background. This tuitional method results a structure which is based on not the goal of algorithms, but on the theoretical basics which we need to know for understanding them. So, let us start it.

7.1 The BB84 protocol

Question: what does *cryptography* mean?

Answer: it is the art of secret communication.

Let us consider the following situation where Alice wants to send a secret message to Bob. In a case like this Alice's message has to be encoded using some kind of cryptographic key. Probably the oldest and simpliest method to make a message secret is the *Caesar code/cipher*, which is a substitutive encryption procedure, where each letter of the alphabet is replaced by another letter. The distance between the position of the original letter and its substitutive is given as a key for the encryption method and its value can have $(1,\ 2,\ 3,\ ....\mathrm{etc})$. In this procedure – in Caesar's time in the history – the primary alphabet was the Latin alphabet and the secondary alphabet was obtained by using this ciphering method. Nowadays, using English alphabet, if "*I love this semester!*" is the message to be encrypted and the value of the letter-transpose (the key) is 2, each letter of the English alphabet will be transposed by 2 places and the encrypted message will be "*k nqxg vjku ugoguvgt!*"[1]. Evidently this is an easy breakable kind of ciphering. However, if we use a random value of transpose for each letter and we use the obtained key only once, we get an unbreakable[2] encryption method called *Vernam cipher* or in other words *One Time Pad* (OTP). (The only disadvantage of this method is that the used key is as long as the text itself.)

So, Alice needs to have two important things:

- One of them is the *algorithm* of the encryption. (For example transposing the letters.)

[1] Provided we do not handle upper and lower cases.
[2] Proved by Shannon.

- The other one is the *key*. For example, using a Caesar code, it is the value of the distance of the transpose.

Nevertheless, we must keep in mind the following – proved – theorem:

the safety of a crypto-system does not depend on the secrecy of the applied algorithm, but depends on the secrecy of the key only. Hence, to find a secret/safe way to share/distribute the encryption key to the parties of the communication is the most important task to solve. This problem can not be solved by using any of the classical methods.

However, we have *a new hope*, because applying one of the several *quantum key distribution* protocols the problem is solvable, with a complete safety.

Let us consider a present-day example to understand how encryption works. So, Alice's message consists of a set of bits, like this one below:

$$0 \quad 1 \quad . \quad . \quad . \quad . \quad 1 \quad 1 \quad . \quad . \quad . \quad . \quad 0 \quad 1$$

In addition to this, she has a randomly generated encryption key consisting a set of bits, too:

$$1 \quad 1 \quad . \quad . \quad . \quad . \quad 1 \quad 0 \quad . \quad . \quad . \quad . \quad 0 \quad 1$$

She can encrypt her message by adding the corresponsive elements of her message and the key, according to *modulo-2*, where

$$0 \oplus 0 = 0,$$
$$0 \oplus 1 = 1,$$
$$1 \oplus 0 = 1,$$
$$1 \oplus 1 = 0.$$

The resulted encrypted message can be seen below:

Original message	0	1	.	.	.	.	1	1	.	.	.	.	0	1
Encryption key	1	1	.	.	.	.	1	0	.	.	.	.	0	1
Encrypted message	1	0	.	.	.	.	0	1	.	.	.	.	0	0

Now then, this encrypted message will be sent to Bob, who needs to have the encription key used by Alice, provided he wants to find out what Alice wanted to communicate. Supposing Bob has the key, it is natural to ask how will he decode the message he got? The answer is very simple: the received (encrypted) message and the key has to be added to obtain the original message:

Encrypted message	1	0	.	.	.	.	0	1	.	.	.	.	0	0
Encryption key	1	1	.	.	.	.	1	0	.	.	.	.	0	1
Original message	0	1	.	.	.	.	1	1	.	.	.	.	0	1

As we can see, it is an unbreakable – or at least a hard breakable – encryption algorithm. However, it has a weakness, namely it can not be garanteed that Alice and Bob are the only persons who have the secret key. In addition to this the presence of an eavesdropper (let us call her Eve) can not be debunked. Let us see, how quantum cryptography help us to eliminate this problem.

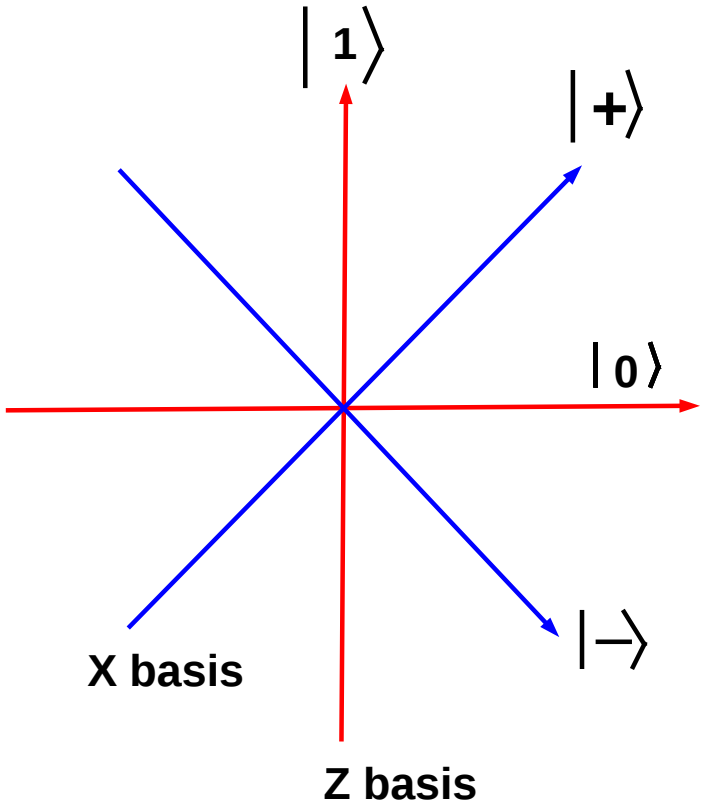
So, the goal of Alice and Bob is to share a common secret key known by themselves only to ensure the secrecy of their communication. Before all, they need to share a quantum and a classical channel (the latter can be anything, even smoke signals). Their common quantum channel is a one way channel from Alice to Bob and the classical channel is a two-way channel. First of all, Alice randomly generate a set of classical bits[3]:

$$0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0$$

After finishing it, Alice encodes each bit of the set into a quantum bit, in the following way: for every qubit, she randomly chooses a preparation basis which can be either the basis spanned by the eigenvectors of σ_z Pauli matrix or another one spanned by the eigenvectors of σ_x Pauli matrix. Let us lebel the former basis as Z basis and the latter one as X basis. As we already know, elements of the Z basis are $\{|0\rangle; |1\rangle\}$ and the two basis vectors of the X basis are $\{|+\rangle; |-\rangle\}$, where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Bearer of the quantum bits used in this protocol can be photons, and their observable polarization (eigen)states play the role of the basis vectors. In this picture, if Z basis is considered as a horizontal polarization basis, then X basis corresponds to the rectilinear polarization basis, where these two bases are the $\pi/2$ rotated of each other.

[3]Naturally, in a real life situation the length of a set like this is much more longer than the presented one.

Now then, depending on the chosen preparation basis, bit 0 can be encoded into either $|0\rangle$ or $|+\rangle$, and similarly, bit 1 can be associated with either $|1\rangle$ or $|-\rangle$. As a visualization, these bases can be imagined in the following graphic way:



In this situation – as in general in any of real life cases – the orientations of the basis elements matter only, or rather we can say the orientations of the axes matter only. It is important to realize and to keep in mind that *if a qubit is in either of the eigenstates of one of these two bases, its state is a superposition on the other basis, and vice versa*. If we keep in mind this fact, we will understand how BB84 quantum key distribution protocol works. Let us consider the figure below, where in the first row, we can see the randomly generated set of classical bits, in the

second row – for each bit – the randomly chosen preparation bases can be seen, and in the third row the prepared quantum bit states are enumerated, where 0-s or 1-s are encoded into each qubit.

	0	1	1	0	1	0	1	1	1	0	1	0
	×	×	+	×	+	+	+	×	×	+	+	×
Quantum bits sent by Alice to Bob:	↗	↖	↑	↗	↑	→	↑	↖	↖	→	↑	↗
	×	+	+	×	×	+	×	+	×	×	+	×
	↗	↑	↑	↗	↖	→	↗	→	↖	↖	↑	↗
	0	1	1	0	1	0	0	0	1	1	1	0
	↗	–	↑	↗	–	→	–	–	↖	–	↑	↗
This is the raw key:	0		1	0		0			1		1	0

From the 4th row, Bob's side is presented. This row containes the measurement bases chosen randomly for each qubit by Bob. What is it all about? After obtaining the qubits sent by Alice, Bob chooses randomly a measurement basis (Z or X) for each qubit he got, and he makes a measurement on every single qubit in the corresponsive basis. After completing his work on measurement, he documents the set of the resulted values. After this, Alice and Bob check the set of their randomly chosen bases (not the results) via a public classical channel, and they keep those elements of their sets of bits where they chose the same basis. The bits which remained form the *raw key*, which – in the presented situation – is the following set:

$$0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0$$

The reason for calling this set *raw* key will be shortly discussed.

One more thing remains to be understood for us to see why this method is better and safer (or rather: absolutely safe) than any of the classical key distributions. First of all let us try to find out what an eavesdropper can do to get the secret key. She (supposing her name is Eve) can catch the quantum bits sent by Alice and before retransmitting them to Bob, she can make a measurement on each quantum bit she caught. And *this is a critical moment*, because she does not know which bases were chosen by Alice and which bases will be chosen by bob (because Alice and Bob will check their bases publicly only *after* Bob obtains the qubits). Hence Eve must select one of the two used bases (Z or X) and make her measurements "blindly" without any knowledge about Alice's or Bob's bases.

And here comes the essence of this protocol: Alice and Bob know their results are necessarily the same for every qubit where they chose the same basis[4]. As we already know these results form the raw key. Now, we say the reason why we call this kind of key a *raw* key. To obtain the "final" key which will be used in the secret communication between Alice and Bob, they need to compare a part of the raw key publicly (immolating the selected part), because they know the following important thing:

if in these cases – where Alice and Bob chose the same basis – the mesurement basis chosen by Eve were not the basis which were selected by Alice and Bob, Eve's measurement results a state which is a *superposition* in the basis chosen by Alice and Bob. Hence Alice and Bob can find different results with a probability of $1/2$ where their results should be the same because of the superposition caused by Eve's measurement (in her wrong chosen basis). *Hence, the presence of an eavesdropper – in our case: Eve – can be detected* because of the non-reversible disturbance introduced into the quantum state by her measurement. The reason for the $1/2$ value of the probability is that even a superposition state can fall into the "good" eigenstate of Bob's basis with a porability of $1/2$. In this case, the presence of Eve remains undetected. Moreover, it is imaginable for Eve to choose the "good" basis with a probablity of $1/2$. In a situation like this her presence can not be detected by comparing Alice's and Bob's results. Fortunately, Alice and Bob work with a raw key whose length is much more longer than the peresented set, hence they can apply a statistical method which makes possible for them to detect the presence of Eve with a complete security. If case of detecting an eavesdropper, they drop the key and create a new one. Summarized: in case of the quantum key distribution, the presence of an eavesdropper can not be kept dark. *This is the reason why quantum cryptography is unbreakable.*

However..... what if Eve decide to copy / clone each quantum bit she caught before retransmitting it to Bob and in this way she can make two clones of every quantum bit sent by Alice. In this situation, having two clones of each quantum bit, she can make a measurement on one of the clones, in basis Z, and can make another measurement on the other clone, in basis X. In a case like this, it is possible for her to circumvent BB84 protocol.

Let us realize there is a crucial point in Eve's process, namely: we have not known if cloning of an unknown quantum state is a possible quantum map or not. This topic is discussed in the following section.

[4]according to the fundamental laws of quantum mechanics, as we already know

7.2 Quantum cloning

First of all, we have to make the meaning of word *quantum cloning* clear. Suppose we would like to build a machine which is able to create a perfect replica of an arbitrary system being in an unknown quantum state. A tool like that seems to be neceassary for certain information processing tasks. Error correction, for instance, could be done using procedures making use of several perfect copies of the original system carrying the information. Such a creation of one or more exact replicas of physical systems in arbitrary (and unknown) quantum states is termed as quantum cloning. The reason for the name, as we shall see, is that the "cloned" system cannot be in fact distinguished from the original one.

It is natural to ask if the laws of quantum mechanics allow us to build such a machine. To put it formally, we consider e.g. a quantum bit in the state. $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$. In addition we need an ancillary sytem which will be the replica. Its initial state can be arbitrary, say $|0\rangle$ w.l.o.g. The desired operation is then

$$|\Psi\rangle|0\rangle \rightarrow |\Psi\rangle|\Psi\rangle. \tag{7.1}$$

Let us first assume that an arbitrary state, say $|\Psi_1\rangle$ can be simply cloned by a unitary operator U:

$$U|\Psi_1\rangle|0\rangle = |\Psi_1\rangle|\Psi_1\rangle. \tag{7.2}$$

If our machine works as we expected, we can continue cloning with another state $|\Psi_2\rangle$. The state of the target qubit is the same before. In this case we get the following states:

$$U|\Psi_2\rangle|0\rangle = |\Psi_2\rangle|\Psi_2\rangle. \tag{7.3}$$

Due to the unitarity, inner product of the left sides of equations 7.2 and 7.3 has to equal the inner product between the right sides of these equations. Hence we obtain the equation below:

$$\langle 0|\langle\Psi_1|U^\dagger U|\Psi_2\rangle|0\rangle = \langle\Psi_1|\Psi_2\rangle^2.$$

After simplifying, we get the following form:

$$\langle\Psi_1|\Psi_2\rangle = \langle\Psi_1|\Psi_2\rangle^2. \tag{7.4}$$

From equation 7.4, it directly follows that

$$\langle\Psi_1|\Psi_2\rangle = \begin{cases} 0 \\ 1 \end{cases}. \tag{7.5}$$

As we can see in equation (7.5), our basic assumption (namely: quantum cloning is a unitary map) leads us results which can be true if and only if we have a total knowledge of the states to be cloned. Obviously, if we knew everyting about these states, we would be able to create them without using any device to clone.

More generally it can be shown that the cloning map in equation (7.1) is not completely positive, so it is not physical. And this holds not only for quantum bits, but also for any kind of quantum systems. This is the *no cloning theorem* of quantum mechanics first pointed out by Żurek [3].

While thus far we have argued that cloning would be a useful operation in information processing, it is easy to see that the fact of its impossibility has also positive implications from practical point of view. For instance it is a basic ingredient of quantum cryptography. If quantum cloning – in the sense of creating perfect replicas of an unknown state – were a possible map, this protocol would be breakable, because – as we saw in the last section – an eavesdropper, after the cloining of the quantum system transmitted between the parties, could achieve measurements on the clones of the qubits sent by Alice to Bob and the quantum key distribution were not secure anymore.

7.3 Conclusion

As we saw, there are two things which makes quantum cryptography unbreakable. On one hand, if Eve makes measurements on the quantum bits she caught, her presence will become visible for the parties of the secret communication. On the other hand, she can even try to make perfect replicas of the transmitted quantum bits to circumvent the protocol, but quantum cloning is forbidden by the fundamental laws of quantum mechanics, hence she does not have any possibility to break the safety of quantum cryptography.

Nevertheless, as usual, a real life situation is never as simple as the presented one. For example, quantum channels are not ideal channels, but more or less they are noisy. In addition to this, although, to make perfect replicas of a quantum system which is in an unknown quantum state is an impossible quantum map – as we saw – , imperfect clones can be done by a device called *quantum cloner*[5]. In case the effect introduced by the cloning process is smaller than the noise caused by the channel, an eavesdropper can pass undetected. Naturally (and fortunately) there are strategies to eliminate this problem, but the discussion of these methods points

[5]This device and some related things will be shortly presented in a future chapter.

beyond the frame of our short course.

From the following chapter we begin to learn about *quantum entanglement*. Besides being a very interesting topic, without this phenomenon, several quantum protocols (e. g. quantum teleportation, dense coding, E91 quantum key distribution protocol, some of the quantum error correction processes, etc.) were not possible.

Chapter 8

Quantum entanglement

Let us get acquinted with the phenomenon of quantum entanglement and outline its aspects which are relevant for our course. We say that the quantum state $|\Psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ of a bipartite system is *separable* if it is a product of states of each subsystem:

$$|\Psi\rangle = |\Psi_1\rangle\,|\Psi_2\rangle\,, \quad |\Psi_1\rangle \in \mathcal{H}_1,\ |\Psi_2\rangle \in \mathcal{H}_2. \tag{8.1}$$

If a state is not separable, it is *entangled*.

The definition can be obviously extended to multipartite systems. And as we shall see later, the entanglement of multipartite systems bears a rich structure.

If a pure state $|\Psi\rangle$ is separable, then all the subsystems are in a pure state. Thus their density operators are projectors. E.g. for

$$\varrho^{(1)} = \mathrm{Tr}_2\,|\Psi\rangle\,\langle\Psi| \tag{8.2}$$

we have

$$\left(\varrho^{(1)}\right)^2 = \varrho. \tag{8.3}$$

As $\mathrm{Tr}\,\varrho = 1$, it implies that

$$\mathrm{Tr}\left(\varrho^{(1)}\right)^2 = 1. \tag{8.4}$$

This holds for all the subsystems if and only if the state is separable.

Considering a bipartite system this leads to a possibility of quantifying entanlgement: the "more mixed" a subsystem is, the more entangled the state is. The mixedness of the state is commonly measured by the von Neumann entropy of the density operator

$$H(\varrho) = -\,\mathrm{Tr}(\varrho \log_2 \varrho), \tag{8.5}$$

which bears a sound information theoretic interpretation. In a d-dimensional system its maximum value is $\log_2 d$, attained by the state

$$\varrho_{\text{CM}} = \frac{1}{d}\hat{I}. \tag{8.6}$$

This is termed as the completely mixed state. It is the only state which produces a uniform distribution of measurement results when measured in any possible basis. For reasons not detailed here the partial traces of a pure bipartite state have the same von Neumann entropy. Hence it is reasonable to say that the *entanglement* of the state is quantified by

$$E\left(|\Psi\rangle\right) = H\left(\text{Tr}_2 |\Psi\rangle\langle\Psi|\right) \tag{8.7}$$

For practical reasons it is worth mentioning that a mathematically simpler quantity can also be used to quantify the mixedness of the state, and thus entanglement, albeit without an operational or direct information theoretic meaning. Its construction stems from the fact that Eq. (8.4) holds if the state is pure. As the diagonal elements of the density matrix describe a probability distributions, for mixed states we have

$$\text{Tr}\,\varrho^2 < 1. \tag{8.8}$$

Hence, the trace of the square of the density matrix is related to the purity of the state in a way. It can be shown that its minimum value is $1/d$ attained by the completely mixed state only. For quantum bits (i.e. $d = 2$ we can thus construct a quantity with in the $[0, 1]$ range (just like the von Neuman entropy):

$$H_{\text{lin}}(\varrho) = 2\left(\frac{1}{2} - \varrho^2\right). \tag{8.9}$$

This is termed as the *linear entropy* of the state. It can be easily verified that the von Neumann entropy is a monotone function of the linear entropy, and so that of its square root. Hence, entanglement can be described also in terms of concurrence

$$C\left(|\Psi\rangle\right) = \sqrt{H_{\text{lin}}\left(\text{Tr}_2 |\Psi\rangle\langle\Psi|\right)}. \tag{8.10}$$

The entanglement in Eq. (8.7) is its monotone function in the same range, it can be evaluated with less effort, but does not admit an operational interpretation.

8.1 Mixed state entanglement

If a multipartite system is in a mixed state, its entanglement properties are far more complex. As for the definition, a mixed state is said to be a separable one,

if it can be constructed as a convex combination of separable pure states or – in other words – it has a form like

$$\varrho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \tag{8.11}$$

in which every $|\Psi_i\rangle$ is separable. Unseparaple mixed states are called entangled.

In many cases it is hard even to decide if a state is separable or entangled at all: obviously in this case the subsystems of a separable state may well be mixed. (Consider the complete mixture of two qubits as an example. It is obviously separable (the density operator being proportional the equal-weight convex combination of the projectos of an arbitrary orthonormal basis, including any product-state basis). Both subsystems are in a completely mixed state though.) Also, while pure-state entaglement is fully characterized by the quantity in Eq. (8.7), for mixed states there are several similar quantities which coincide for pure states but they have different operational meanings otherwise.

One of them is *entanglement of formation* defined as follows:

$$E(\varrho) = \inf_{\substack{(p_k,\ |\psi_k\rangle \text{ separable}) \\ \sum_k p_k |\psi_k\rangle\langle\psi_k| = \varrho}} \sum_k p_k E(|\psi_k\rangle), \tag{8.12}$$

that is, the infimum of the average of the entaglements of all the constituent pure states over all of its pure-state decompositions. As we deal with finite dimensional states, the infimum can be understood as minimum.

A similar quantity can be defined via concurrence:

$$C(\varrho) = \inf_{\substack{(p_k,\ |\psi_k\rangle \text{ separable}) \\ \sum_k p_k |\psi_k\rangle\langle\psi_k| = \varrho}} \sum_k p_k C(|\psi_k\rangle), \tag{8.13}$$

It can be shown that entanglement of formation is its monotone function, and in the special case of two qubits it can be calculated analytically. This is the celebrated Wootters formula which is very broadly used in the literature, including our work. Hence we describe it in what follows. For the detailed derivations we refer to the original papers.

8.2 The Wootters formula

In order to calculate the concurrence of a two-qubit state ϱ, first we define the Wootters tilde operation:

$$\tilde{\varrho} = (\sigma_y \otimes \sigma_y)\varrho^*(\sigma_y \otimes \sigma_y), \tag{8.14}$$

where $*$ means the complex conjugation (or, otherwise speaking, the transpose) of the density matrix in a product state basis, whereas σ_y is the second Pauli-operator.

Next the spectrum of the Hermitian operator has to be determined

$$\sqrt{\sqrt{\varrho}\tilde{\varrho}\sqrt{\varrho}}. \tag{8.15}$$

Its eigenvalues λ are in fact the square roots of the eigenvaluses of the (non-Hermitian) operator

$$\varrho\tilde{\varrho}. \tag{8.16}$$

Let us put the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ to decreasing order. The concurrence then reads

$$C(\varrho) = max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \tag{8.17}$$

We shall employ this latter formula when calculating concurrence in some future chapters.

8.3 Entanglement of multi-qubit systems

It may be an interesting question how can be featured the entanglement of two chosen quantum bits in a system consisting of N quantum bits, if the whole system is in a pure state. As an illustration, let us consider the following specific example: we have three quantum bits in a state which is called Greenberger-Horne-Zeilinger (GHZ) state:

$$|\Psi_{\mathrm{GHZ}}\rangle_{123} = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle). \tag{8.18}$$

In this case, the state of the first two quantum bits is describes by the density operator

$$\varrho_{12} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|). \tag{8.19}$$

This state is obviously entangled. In fact, all the subsystems are in a completely mixed state.

When considering any of the two qubits (e.g. the first two, any of them can be choosen for symmetry reasons), however, using the formula in (8.17), for this density operator, we get $C(\varrho_{12}) = 0$. This means that state (8.19) is a separable state or, in other words, the first two quantum bits are not entangled with each other as a pair.

It means that in the present entangled state there is no qubit-pair entanglement whatsoever. Indeed, after carrying out a measurement on the third quantum bit in the basis $|0\rangle$, $|1\rangle$, the state of the first two quantum bits will be either $|00\rangle$ or $|11\rangle$ with equal probability. This means that the bipartite state can be constructed as a convex combination of separable pure states.

So the entanglement of the first two quantum bits can be juggled away by achieving measurement on the third one. Due to the symmetry of the state, this holds true of the case of any pair of quantum bits in this state. In the GHZ state (8.18), the state of any quantum bit pair can be separated. The whole system is an entagled state, after all! State (8.18) is not separable. This can be seen, if we choose one of the three quantum bits, its state, according to (8.6), is a maximally mixed state, that is, the chosen quantum bit is maximally entangled with the subsystem of the other two quantum bits.

Hence, the entanglement present in this state is *genuine threepartite*. Interestingly, it can be converted to maximal bipartite entanglement though. Carrying out a properly chosen measurement on one of the quantum bits, we can make the state of the system of other two quantum bits maximally entangled. Indeed, if the eigenvectors of the measurement are now $\frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$, we get the maximally entangled states $\frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ with equal probability. Both states are maximally entangled bipartite states. If we are aware of the measurement result, we know which state we have obtained, so we can use it, e.g. for teleportation[1]. In this system, its tripartite entanglement can be completly converted into a bipartite entanglement.

8.4 Monogamy of entanglement

Note that a maximally entangled state of two quantum bits is necessarily a pure state. Hence two quantum bits cannot be entangled with any other system. This means that (unlike classical correlations) quantum entanglement has a property which is called *monogamy*: pairwise entanglement of two subsystems limits the

[1] more on this later

entanglement with the other subsystems.

As for a quantitative description of monogamy, we introduce another quantity which we will use in our work. This is the *tangle* denoted by τ, which is the linear entropy in Eq. (8.9) of a given subsystem, which, for qubits can also expressed as

$$\tau_k = 4\mathrm{det}\varrho(k).$$

This is the so-called one-tangle, characterizing the entanglement between the given qubit and the rest of the total system which is in a pure state. In case we have two qubits in a pure state, the tangle relating to one of them equals the square of the concurrence. Let us consider a system consisting of many qubits and suppose the system is in a pure state. Checking concurrences of the qubit pairs in the system, we get that Coffmann-Kundu-Wootters (CKW) inequalities [2] are satisfied:

$$\tau_k \geq \sum_{l \neq k} C_{k,l}^2 \tag{8.20}$$

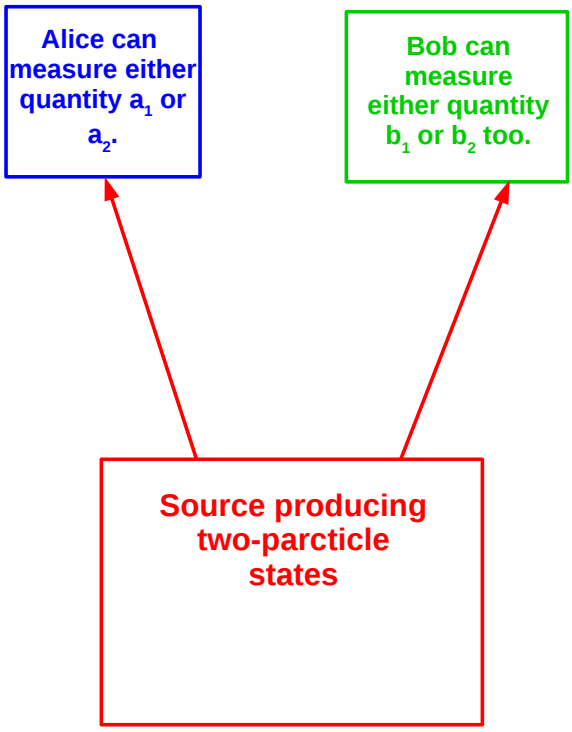
This formula can be interpreted in the following way: the entanglement measured in tangle between the *k*-th qubit and the rest of the total system gives an upper bound for the concurrence calculated between the *k*-th qubit and another arbitrary qubit in the system. If these inequalities are saturated, the bipartite entanglement is maximal.

The CKW inequalities had been originally formulated as a conjecture, but they were later proven. Their saturation reflects that the bipartite entanglement is in a way maximal in the system.

8.5 Bell inequalities

Since entangled states can have stronger correlations than any of the classical correlations can be, as we already mentioned it, they are a valuable resource in quantum communication. Before getting on the protocols based on the phenomenon of quantum entanglement, let us get to know Bell's inequalities to understand what nonclassical correlations mean. As we know, in quantum mechanics, an observable physical quantity does not have value until we measure it. However, there were (and are) theories according to which – unlike in quantum mechanics – , observables have values, even we have not carried on a measurement on them. Since they depend on some "hidden variables", we do not know their actual values, because we do not know anything about these hidden variables. Bell's inequalities show that in case of very general conditions, hidden variable theories

(more precisely: *local* hidden variables theories) yield predictions which conflict with quantum mechanics and – what is more – these can be experimentally tested. Spoiler: experimetal tests support quantum mechanics and reject the locality of our world. In essence – as we shall see – Bell's inequalities make a philosophical debate testable experimentally. And this was / is the biggest merit of Bell and his inequalities. In case of Bell's inequalities, we can see two observers (see the figure below), namely: Alice and Bob (who else?). Between them, there is a source producing two-particle states. One of these particles is sent to Alice and the other one to Bob. On her particle, Alice can measure one of two quantities, a_1 and a_2. The resulted values of these observable quantities can be either 1 or -1. Similarly, Bob can measure either b_1 or b_2 , and the measured values can equal either 1 or -1, too. The essence of this idea is to execute this gedankenexperiment many times, and use the results to calculate the quantities $\langle a_i b_j \rangle$.



First of all, let us see how a hidden-variable theory would describe this case.

The source produces regulation sets which go with the particles. For instance, one of these regulation set can say, in case Alice measures a_1, she will get 1, and measuring a_2 , she gets -1, furthermore if Bob measures b_1 he will get -1, and in case of measuring a_2 he gets -1. We do not know which regulation set will be produced by the source, hence the regulation set is our hidden variable. This kind of a hidden-variable theory are called *local* theory, because the regulations to Alice's particle do not depend on Bob's decision on the quantity to be measured. That is, the regulation set does not say anything like, measuring a_1, Alice will obtain 1 if Bob measures b_1 and she gets -1 if Bob measures b_2. We will consider local theories only.

In a situation like this, a hidden variable can be the state of the source. If we know the source is in some state m, results of the measurments can be prognosticated. Obviously there are 16 possibilities:

value of m	a_1	a_2	b_1	b_2
1	-1	-1	-1	-1
2	-1	-1	-1	1
3	-1	-1	1	-1
4	-1	-1	1	1
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
16	1	1	1	1

Supposing we do not have access to the source, we assume $P(m)$ is a distribution function of the states of the source, or equivalently, a certain number foursome $(a_1,\, a_2,\, b_1,\, b_2)$ can appear with a given probability. This means there is a $P(a_1,\, a_2,\, b_1,\, b_2)$ distribution function. From this it follows that the expectation value $\langle a_1 b_1 \rangle$ can be calculated as

$$\langle a_1 b_1 \rangle = \sum_{a_1=1}^{-1} \sum_{a_2=1}^{-1} \sum_{b_1=1}^{-1} \sum_{b_2=1}^{-1} a_1 b_2 \, P(a_1,\, a_2,\, b_1,\, b_2).$$

There can be possible 4 pieces of correlation functions like this, namely: $\langle a_1 b_1 \rangle$, $\langle a_1 b_2 \rangle$, $\langle a_2 b_1 \rangle$, $\langle a_2 b_2 \rangle$. According to calculations (not detailed here), the following expression yields the biggest value which can be reached by this kind of (classical) correlations:

$$S = \langle a_1 b_1 \rangle + \langle a_1 b_2 \rangle + \langle a_2 b_1 \rangle - \langle a_2 b_2 \rangle =$$
$$= \sum_{a_1=1}^{-1} \sum_{a_2=1}^{-1} \sum_{b_1=1}^{-1} \sum_{b_2=1}^{-1} [a_1(b_1 + b_2) + a_2(b_1 - b_2)]\, P(a_1,\, a_2,\, b_1,\, b_2).$$

Let us call the expression in brakets multiplying the probability distribution X. As it can be seen

$$X = \begin{cases} a_1(b_1 + b_2), & \text{if} \quad b_1 = b_2 \\ a_2(b_1 - b_2), & \text{if} \quad b_1 \neq b_2. \end{cases}$$

In both cases $|X| = 2$, hence

$$|S| \leq 2 \sum_{a_1=1}^{-1} \sum_{a_2=1}^{-1} \sum_{b_1=1}^{-1} \sum_{b_2=1}^{-1} P(a_1,\, a_2,\, b_1,\, b_2) = 2.$$

Expression $|S| \leq 2$ is Bell's inequality. Naturally, similar inequalities can be derived simply by interchanging a_1 and a_2 , b_1 and b_2, or both.

Now, supposing we are measuring the spins of two half-spin particles, we describe the same experiment using the apparatus of quantum mechanics. (a_1 and a_2 (just like b_1 and b_2) can be considered as measurements of the spin (or polarization of a photon) along two different directions.)

Having a quantum source, let us suppose that

$$a_1 = \sigma_{x_a} \quad a_2 = \sigma_{y_a}$$
$$b_1 = \sigma_{x_b} \quad b_2 = \sigma_{y_b},$$

and that the source emits particles in a state which is a maximally entangled pure (bipartite) state (multiplied with a phase factor), or in other words, one of the four Bell states:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \big(|00\rangle + e^{i\frac{\pi}{4}} |11\rangle \big),$$

where

$$\sigma_x|0\rangle = |1\rangle \quad \sigma_y|0\rangle = i|1\rangle$$
$$\sigma_x|1\rangle = |0\rangle \quad \sigma_y|1\rangle = -i|0\rangle.$$

In this case

$$\langle a_1 b_1\rangle = \langle\Psi|\sigma_{x_a}\otimes\sigma_{x_b}|\Psi\rangle =$$
$$= \frac{1}{2}\Big[\big(\langle 00| + e^{-i\frac{\pi}{4}}\langle 11|\big)\sigma_{x_a}\otimes\sigma_{x_b}\big(|00\rangle + e^{i\frac{\pi}{4}}|11\rangle\big)\Big] =$$
$$= \frac{1}{2}\big(\langle 00| + e^{-i\frac{\pi}{4}}\langle 11|\big)\big(|11\rangle + e^{i\frac{\pi}{4}}|00\rangle\big) = \frac{1}{2}\big(e^{i\frac{\pi}{4}} + e^{-i\frac{\pi}{4}}\big) =$$
$$= \frac{1}{2}2\cos\frac{\pi}{4} = \frac{1}{2}2\frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{2}.$$

Since

$$\langle a_1 b_1\rangle = \langle a_1 b_2\rangle = \langle a_2 b_1\rangle = \frac{\sqrt{2}}{2}$$

and

$$\langle a_2 b_2\rangle = -\frac{\sqrt{2}}{2},$$

it directly follows that

$$S = \langle a_1 b_1\rangle + \langle a_1 b_2\rangle + \langle a_2 b_1\rangle - \langle a_2 b_2\rangle = 4\frac{\sqrt{2}}{2} = 2\sqrt{2} \geq 2,$$

that is quantum mechanics violates Bell's inequality. This fact has three consequences:

- Quantum mechanics – that is our world – can not be described by a local, hidden variable theory. From this, it follows that our world is nonlocal(!).
- In the local hidden variable theory, correlations came from a joint probability distribution function.

- Quantum mechanics can create stronger correlations than classical systems can.

Chapter 9

Some simple applications of entanglement

9.1 Dense coding

Using this method, Bob can send Alice (or vice versa) two bits of classical information by transmitting only one qubit. This protocol is based – what a surprise – on quantum entanglement.

In the beginning, Alice and Bob share an entangled pair of qubits whose state is one of the four Bell states:

$$|\Phi_-\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|01\rangle_{AB} - |10\rangle_{AB}\big), \tag{9.1}$$

where indeces A and B refer to the name of the person who has the given member of the pair. As we know, the four Bell states form an orthonormal basis in the four dimensional Hilbert space, and they have the following shapes:

$$|\Phi_\pm\rangle = \frac{1}{\sqrt{2}}\big(|01\rangle \pm |10\rangle\big)$$
$$|\Psi_\pm\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle \pm |11\rangle\big).$$

Suppose Bob performs one of the four preconcerted unitary operations on the

qubit he has. The four possible operataions are enumerated below.

- she applies operator $\hat{I}$ on her qubit: or in other words, she does nothing
- she applies operator σ_x
- she applies operator σ_y
- she applies operator σ_z

Let us recall the shapes of the three Puali matrices and the unit operator, then try to find out their effect on Alice's single qubit, finally calculate the resulted two-particle states. So, the shape of the three Pauli matrices and the unit operator can be seen below:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \hat{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Now, let us see their effect on a single quantum bit whose state can be either $|0\rangle$ or $|1\rangle$:

$$\sigma_x|0\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle$$

$$\sigma_x|1\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle.$$

The other four results can be derived similarly:

$$\sigma_y|0\rangle = i|1\rangle, \quad \sigma_y|1\rangle = -i|0\rangle, \quad \sigma_z|0\rangle = |0\rangle, \quad \sigma_z|1\rangle = -|1\rangle$$

Finally, we are interested in the two-particle states resulted by the four different operations of Bob. Let us see how will the initial entangled state in 9.1 be altered by Bob's operations. In case he works with operator $\hat{I}$, it is evident the state remains the same:

$$|\Phi_-\rangle = \frac{1}{\sqrt{2}}\big(|01\rangle - |10\rangle\big),$$

where indeces A and B are omitted now. If he apply σ_x on his qubit the resulted entangled state is

$$\frac{1}{\sqrt{2}}\big(|00\rangle - |11\rangle\big).$$

In case of applying σ_y, they get the following two-particle state:

$$-i\frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big).$$

Finally, in case of applying operator σ_z on Bob's single qubit, the common state of the two entangled qubits they share will be changed into the following state:

$$-\frac{1}{\sqrt{2}}\big(|01\rangle + |10\rangle\big).$$

Let us realize the four two-particle states which can be resulted by Bob's operation are actually the elements of the orthonormal Bell basis. The only difference between Bell states (which form the Bell basis) and the state we can obtain is two scalar factors $(-i,\, -1)$. Hence any of Bob's operations results a state which is one of the elements of a four dimensional, orthogonal basis.

Knowing this very important fact, we can turn to the second step of the protocol: *after* performing his operation, Bob sends his qubit to Alice, hence from the moment she gets Bob's qubit, Alice has the whole system, hence setting four othogonal detectors – correspondently the four possible states – she can find out the resulted state and the corresponsive operation.

There is a crucial point here which has to be realized, namely: Alice's measurement selects one of the *four* possibilities which means *two bits* of classical information, though Bob sent *one* piece of *quantum bit*..... and this is the reason for the name of this protocol, namely: *dense* coding.

9.2 Quantum teleportation

In this protocol, Alice has a qubit denoted by A_1 in a state

$$|\Psi\rangle_{A_1} = \alpha|0\rangle + \beta|1\rangle,$$

and she would like to transfer this state onto Bob's quantum bit denoted by B. It is worth notify that Alice does not even need to know anything about the state of her qubit. The question is what she can do to reach her goal. As a first idea one can suggest Alice to measure the state of her qubit and to transmit the resulted classical information to Bob. Unfortunately this strategy would not work, because - as we know - the obtained information is not enough to reconstruct an arbitrary, unknown state[1].

In the method of quantum teleportation, Alice and Bob share an entangled pair, A_2, B in one of the Bell states:

$$|\Phi_-\rangle_{A_2B} = \frac{1}{\sqrt{2}}\big(|01\rangle_{A_2B} - |10\rangle_{A_2B}\big).$$

The whole state of their three qubits, the one whose state is to be teleported, and the entangled pair can be expressed as a tensor product:

$$|\Psi\rangle_{A_1}|\Phi_-\rangle_{A_2B} = \frac{1}{\sqrt{2}}\big(\alpha|0\rangle_{A_1} + \beta|1\rangle_{A_1}\big)\big(|01\rangle_{A_2B} - |10\rangle_{A_2B}\big) =$$

$$= \frac{1}{\sqrt{2}}\big(\alpha|00\rangle_{A_1A_2}|1\rangle_B - \alpha|01\rangle_{A_1A_2}|0\rangle_B + \beta|10\rangle_{A_1A_2}|1\rangle_B - \beta|11\rangle_{A_1A_2}|0\rangle_B\big)$$

Instead of using a product basis, let us rewrite this expression into another shape via the elements of the Bell basis spanned by the four basis vectors below:

$$|\Phi_\pm\rangle = \frac{1}{\sqrt{2}}\big(|01\rangle \pm |10\rangle\big)$$

$$|\Psi_\pm\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle \pm |11\rangle\big).$$

From these elements, it directly follows that

$$\sqrt{2}|\Phi_\pm\rangle = |01\rangle \pm |10\rangle$$

$$\sqrt{2}|\Psi_\pm\rangle = |00\rangle \pm |11\rangle.$$

Using these equations, after a trivial calculation we obtain that

[1] We know this is caused by the effect of the measurement on the state vector.

$$\frac{1}{\sqrt{2}}\big(|\Phi_+\rangle + |\Phi_-\rangle\big) = |01\rangle$$
$$\frac{1}{\sqrt{2}}\big(|\Phi_+\rangle - |\Phi_-\rangle\big) = |10\rangle$$
$$\frac{1}{\sqrt{2}}\big(|\Psi_+\rangle + |\Psi_-\rangle\big) = |00\rangle$$
$$\frac{1}{\sqrt{2}}\big(|\Psi_+\rangle - |\Psi_-\rangle\big) = |11\rangle.$$

If we use these results, we can continue rewriting the initial expression of the whole state, namely

$$|\Psi\rangle_{A_1}|\Phi_-\rangle_{A_2B} = \frac{1}{2}\Big[\alpha\big(|\Psi_+\rangle + |\Psi_-\rangle\big)|1\rangle_B - \alpha\big(|\Phi_+\rangle + |\Phi_-\rangle\big)|0\rangle_B +$$
$$\beta\big(|\Phi_+\rangle - |\Phi_-\rangle\big)|1\rangle_B - \beta\big(|\Psi_+\rangle - |\Psi_-\rangle\big)|0\rangle_B\Big] =$$
$$\frac{1}{2}\Big[|\Psi_+\rangle_{A_1A_2}\big(\alpha|1\rangle_B - \beta|0\rangle_B\big) + |\Psi_-\rangle_{A_1A_2}\big(\alpha|1\rangle_B + \beta|0\rangle_B\big) +$$
$$|\Phi_+\rangle_{A_1A_2}\big(-\alpha|0\rangle_B + \beta|1\rangle_B\big) + |\Phi_-\rangle_{A_1A_2}\big(-\alpha|0\rangle_B - \beta|1\rangle_B\big)\Big]$$

The key element of this process can be seen in the last two rows. What is it all about? We can see Alice has two quantum bits, A_1 and A_2. What if she makes a Bell measurement on the system of these two qubits (which is the subsystem of the whole threepartite qubit system)? Let us remember, in case of a Bell measurement on a bipartite system like Alice has, one of the four Bell states will be resulted. As we can see in the last two rows, after making her measurement, the (eigen)state of Alice's subsystem can be one of the well known Bell states. Nevertheless, what is more, let us focus on the state of Bob's qubit after Alice's measurement, say, what if the resulted state of Alice's qubits is $|\Phi_+\rangle_{A_1A_2}$? As we can see, in this case the state of Bob's qubit will be $-\alpha|0\rangle_B + \beta|1\rangle_B$, because the whole threepartite system will get to the state $|\Phi_+\rangle_{A_1A_2}\big(-\alpha|0\rangle_B + \beta|1\rangle_B\big)$ by Alice's measurement.

Why is it so important? To answer this question we need to realize the fact that if Bob knew Alice's result ($|\Phi_+\rangle_{A_1A_2}$), he would know what to do to get the state Alice wanted to teleport to him. Four instance, in the discussed case, he would make a unitary transformation on his qubit state. The question is what kind of

transformation should be done by Bob? The answer is nearly trivial: the respective operator of the transformation is $-\sigma_z$. Let us try to find out if it is true or not: the matrix which represents $-\sigma_z$ is

$$-\sigma_z = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Once again: in case Alice's measurement results the state $|\Phi_+\rangle_{A_1A_2}$, Bob's qubit will be in $-\alpha|0\rangle_B + \beta|1\rangle_B$. Let us see the effect of $-\sigma_z$ on this state:

$$-\sigma_z\big(-\alpha|0\rangle_B + \beta|1\rangle_B\big) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -\alpha \\ \beta \end{pmatrix} =$$

$$= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha|0\rangle_B + \beta|1\rangle,$$

which means the state of Bob's qubit after making the transformation is exactly the same like the state which was wanted to teleport by Alice. All she had to do was to send the result of her Bell measurement to Bob.

This is almost exactly the case when any of the three other possible outcomes is resulted by Alice's measurment. The only difference is the kind of the transformation which has to be applied by Bob on his quantum bit to get the state to be teleported. These "result – transformation" pairs are scheduled in the table below:

result of Alice's Bell measurement	$\|\Phi_+\rangle$	$\|\Phi_-\rangle$	$\|\Psi_+\rangle$	$\|\Psi_-\rangle$
Bob's transformation	$-\sigma_z$	$-\hat{I}$ (or $\hat{I}$)	$-\sigma_x\sigma_z$	σ_x

Let us summarize the discussed process of quantum teleportation. In the beginning the two parties share an entangled state consisting of two qubits. On Alice's side there is another qubit whose state is wanted to teleport (or in other words to transfer onto Bob's qubit) by Alice. To reach her goal, Alice makes a Bell measurement on her two qubits and then she messages her result to Bob via a classical channel. Knowing Alice's result, Bob knows what kind of transformation has to be applied on his quantum bit to transfer / teleport the state of qubit A_1 onto his quantum bit.

There is a very important fact which we have to realize: this protocol is *not* copying / cloning, because the quantum bit – A_1 – which was in the teleported state initially, will be in a totally different state after applying the protocol.

And another interesting remark: if we do not know the state to be cloned, we can not clone it (if we know the state, we can prepare it and there is no need to clone), on the other hand, we can teleport any of the states, even in the case we do not know anything about it.

9.3 E91 quantum key distribution protocol

In 1991 Artur Ekert suggested a protocol based on shared entanglement instead of sending particles via a quantum channel (as we saw it in the case of BB84 protocol). Let us suppose a source sending entangled quantum bit pairs to Alice and Bob. One member of each qubit pair is sent to Alice and the other one is sent to Bob. Each qubit pair is in a singlet state which is the same in both X basis and Y basis. We already know the elements of X basis, namely: $\{|+\rangle; |-\rangle\}$. Nevertheless, also the basis vectors of Y were mentioned in equation 2.203, where we wrote the eigenvectors of σ_y Pauli matrix. These eigenvectors, $|\pm\rangle_y = \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$, where

$$|+\rangle_y = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix} \quad \text{and} \quad |-\rangle_y = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix},$$

span the Y basis. As we said, the singlet state of the qubit pairs has equivalent expressions on both X and Y basis:

$$\frac{1}{\sqrt{2}}\big(|+\rangle_A|-\rangle_B - |-\rangle_A|+\rangle_B\big) = \frac{1}{\sqrt{2}}\big(|+\rangle_{y_A}|-\rangle_{y_B} - |-\rangle_{y_A}|+\rangle_{y_B}\big)$$

because on X basis, it has the following shape:

$$\frac{1}{\sqrt{2}}(|+\rangle|-\rangle - |-\rangle|+\rangle) = \frac{1}{\sqrt{2}}\left(\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix} - \begin{pmatrix}1\\-1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}\right) =$$

$$= \frac{1}{\sqrt{2}}\left(\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix} - \begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}\right) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\-2\\2\\0\end{pmatrix},$$

and then on Y basis, its form is the expression below:

$$\frac{1}{\sqrt{2}}(|+\rangle_y|-\rangle_y - |-\rangle_y|+\rangle_y) = \frac{1}{\sqrt{2}}\left(\begin{pmatrix}1\\ i\end{pmatrix}\begin{pmatrix}1\\ -i\end{pmatrix} - \begin{pmatrix}1\\ -i\end{pmatrix}\begin{pmatrix}1\\ i\end{pmatrix} \right) =$$

$$= \frac{1}{\sqrt{2}}\left(\begin{pmatrix}1\\ -i\\ i\\ 1\end{pmatrix} - \begin{pmatrix}1\\ i\\ -i\\ 1\end{pmatrix} \right) = i\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ -2\\ 2\\ 0\end{pmatrix}.$$

We can see the singlet state is the same in both bases (as we know a state and itself multiplied by i are the same states). The point of the protocol is that Alice and Bob autonomously choose a measurement basis from between X and Y randomly and make a measurement in the selected bases. After this, they communicate each other the basis they chose (but not the result). In case they chose the same basis, the results of their measurements are perfectly anti-correlated. Hence the presence of an eavesdropper can be detected in exactly the way as we saw in the case of BB84 protocol: they need to check a part of the (raw) key. Nevertheless, there is an important difference, namely: they can use their measurement results obtained in the cases when they chose different bases, hence they can test if a Bell inequality is violated or not. In case Eve had taken over the source and were sending particles in definite states to Alice and Bob, for instance a $|+\rangle$ to Alice and a $|-\rangle$ to Bob, then the Bell inequality would not be violated, and Alice and Bob would detect her.

9.4 Quantum error correction

In this section, we study how quantum entanglement is used in error correction processes. Due to the noisy quantum channels[2], error correction is required in the field of quantum communication protocols. First of all, we have to clarify the reason why we use the quantum entanglement as a tool in error correction, instead of using redundancy as we saw in the case of classical information systems. The answer is very trivial, if we recall what we learned about quantum cloning. Being an impossible quantum map, quantum cloning can not be used to create one or several clones of a qubit which is in an unknown quantum state, hence redundancy can not be applied if we handle quantum bits.

[2]Naturally there are several other phenomena which can cause errors in quantum information protocols, for instance faulty gates, measurements, preparations

Fortunately, using a highly entangled state of several qubits, we have a possibility to spread the information carried by a single qubit onto the entangled state. Peter Shor suggested a method where the information of one qubit is stored in the highly entangled state of nine qubits.

Classical error correcting codes generally use a syndrome measurement to detect which error appears in an encoded state. After this, applying a corresponsive operation based on the detected syndrome we can correct the error. Also in quantum error correction protocols syndrome measurements are used. Making a multi-qubit measurement, we do not disturb the information stored in the encoded state, however we retrieve information about the error. Besides diagnosing if a qubit has been corrupted, we can also find out which qubit was affected, if we use syndrome measurements. In addition, the outcome of a syndrome measurement bespeaks us the way how the qubit was corrupted (bit flip, sign flip or both).

It is natural to ask: how can the effect of the noise be one of a few possibilities, when the noise itself can be arbitrary? In most of the codes, this effect can be either a bit flip, or a sign flip, or both (corresponding to the Pauli matrices X, Z, and Y). It is important to keep in mind that syndrome-measurements are projective measurements (just like any of the quantum measurements (except when POVM is made, but this kind of measurement is not discussed in this semester)), hence though the noise can be arbitrary, it can be built as a superposition of basis operations. The syndrome-measurement makes the qubit to choose one of the Pauli errors to have happened (it is said to be "Pauli error" because of the kinds of errors, for instance bit flip, sign flip, or both), and the syndrome bespeaks us which error was choosen. Hence we can apply the same Pauli operator on the corrupted qubit and in this way the effect of the error can be reverted. Unfortunately syndrome-measurements tell nothing about the value stored in the logical qubit.

9.4.1 The bit flip code

As we know, repetition code can not be used in case of a quantum channel, because of the no-cloning theorem. Fortunately there is a solution of this problem, namely the three-qubit bit flip code. In this method syndrome-measurement and entanglement are applied and this "protocol" is as effective as the repitition code was in classical cases.

Let suppose we want to transmit a qubit $|\psi\rangle$ through a noisy channel C. We know that this channel either flips the state of the qubit, with probability p, or leaves it untouched. Hence, if we have a general qubit ρ as an input state of the channel C,

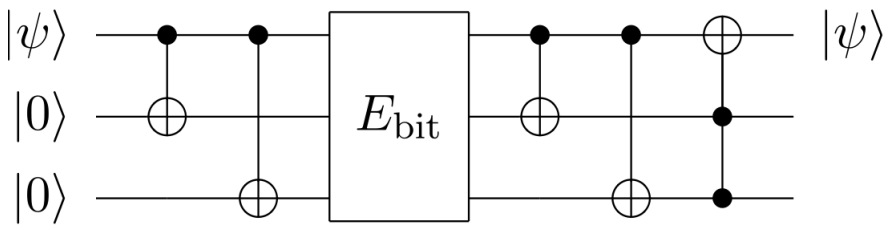


Figure 9.1: Quantum circuit of the bit flip code

the action of the channel on the qubit can be written as

$$C(\rho) = (1-p)\rho + p\, X\rho X.$$

In our case, we want to transmit the following state: $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Without applying any error correcting protocol, the state will be correctly transmitted with a probability of $1-p$. Naturally this number can be improved, if we encode the state into a greater number of qubits. In this way we can detect and correct errors in the corresponding logical qubits.

In the case of the simple three-qubit repetition code, the encoding means the following mappings:

$$|0\rangle \to |0_L\rangle \equiv |000\rangle \quad \text{and} \quad |1\rangle \to |1_L\rangle \equiv |111\rangle.$$

The input state $|\psi\rangle$ is encoded into the state $|\psi'\rangle = \alpha|000\rangle + \beta|111\rangle$. This mapping can be realized in case we apply two CNOT gates to entangle the system with two ancillary qubits initialized in the state $|0\rangle$. Hence, the encoded state $|\psi'\rangle$ is sent through the noisy channel.

The channel can have an effect on $|\psi'\rangle$, because it can flip some subset of the qubits of the state. There is no qubit flip with a probability of $(1-p)^3$, one of the qubits is flipped with a probability of $3p(1-p)^2$, two of them are flipped with a probability of $3p^2(1-p)$ and all of them are flipped with a probability of p^3. Let us realize that there is an additional assumption about the channel, namely: we assume that the effects of C on each qubit of $|\psi'\rangle$ are independent and equal. Thus our problem is how to detect and correct such errors, without corrupting the state to be transmitted.

For the sake of simplicity, let us suppose that p is very small, thus the probability of more than one qubit is flipped is negligible. In this case we can detect if a qubit was flipped, without also querying for the values being transmitted, by asking whether one of the qubits differs from the others. This amounts to performing a measurement with four different outcomes, corresponding to the following four projective measurements:

$$P_0 = |000\rangle\langle 000| + |111\rangle\langle 111|$$
$$P_1 = |100\rangle\langle 100| + |011\rangle\langle 011|$$
$$P_2 = |010\rangle\langle 010| + |101\rangle\langle 101|$$
$$P_3 = |001\rangle\langle 001| + |110\rangle\langle 110|$$

We can do it, if we first measure Z_1Z_2 and then Z_2Z_3. This tells us which qubits are different from which others but does not yield any information of the state of the qubits. In case we obtain the outcome which corresponds to P_0, there is no need to apply any of the corrections. On the other hand, in case of obtaining an outcome corresponding to P_i, we have to apply the Pauli X gate on the i-th qubit. This procedure can be expressed by the following map (which has to be applied on the output of the channel):

$$C_{corr}(\rho) = P_0\,\rho\,P_0 + \sum_{i=1}^{3} X_iP_i\,\rho\,P_iX_i.$$

It is important to know that this procedure does not work properly, if there are more the one qubit flipped by the channel.

9.4.2 The sign flip code

Although in classical case bit flip can be the only kind of error, using a quantum computer we have to face another kind of error, namely the sign flip. This means that the relative sign between $|0\rangle$ and $|1\rangle$ can become inverted. For example a state like this $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ can change into the following state: $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, which is a sign flip.

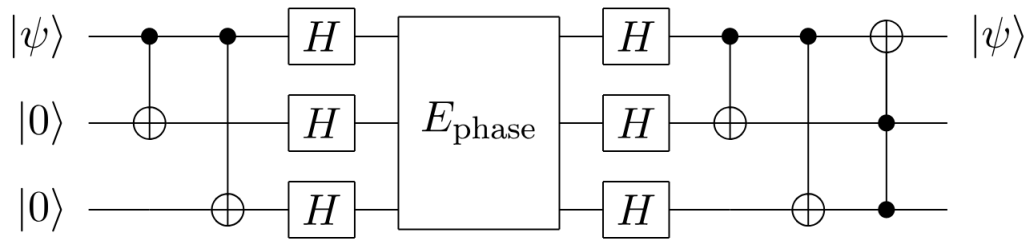


Figure 9.2: Quantum circuit of the phase flip code

The original state to be transmitted is $|\psi\rangle = \alpha|+\rangle + \beta|-\rangle$, which will be altered into the following state:

$$|\psi'\rangle = \alpha|+++\rangle + \beta|---\rangle.$$

Expressing a state on the Hadamard base, a sign flip becomes bit flip and vice versa. In case C_{phase} is a quantum channel causing at most one bit flip, the bit flip code can recover $|\psi\rangle$ by transforming it into the Hadamard basis before and after transmission through C_{phase}.

9.4.3 The Shor code

A noisy channel can cause sign flip, bit flip or both. Using the Shor code both types of errors can be corrected. Actually, arbitrary single-qubit errors can be fixed by applying the Shor code. Let us suppose C is a quantum channel, which can corrupt a quantum bit in an arbitrary way. In our figure, the 1st, 4th and 7th are for the sign flip code and the three group of qubits (1,2,3), (4,5,6), (7,8,9) are designed for the bit flip code. Using the Shor code, we transform the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ into the product state of nine qubits $|\psi'\rangle = \alpha|0_S\rangle + \beta|1_S\rangle$, where

$$|0_S\rangle = \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle)$$

and

$$|1_S\rangle = \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle).$$

In case one of the qubits has a bit flip error, the syndrome analysis is performed on each set of states (1,2,3), (4,5,6), (7,8,9), and the error will be fixed.

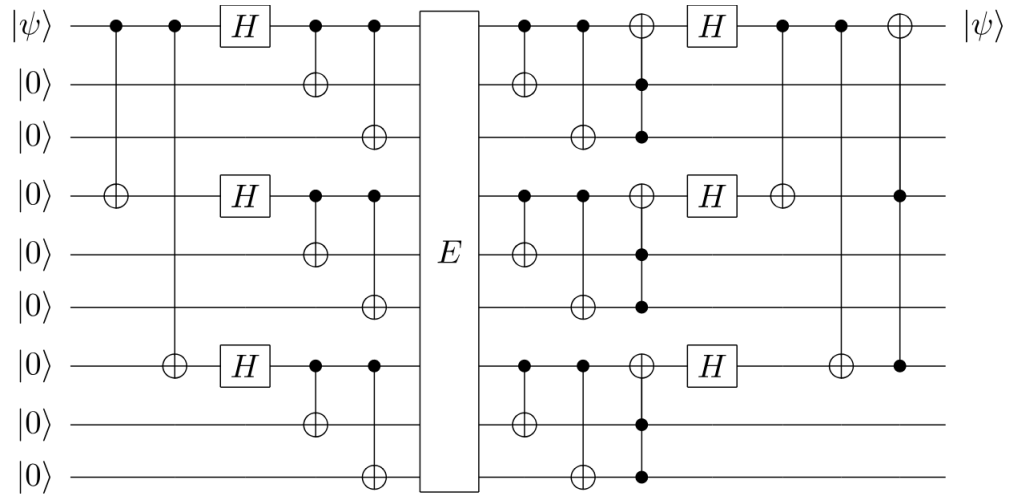


Figure 9.3: Quantum circuit of Shor code

If the three bit flip groups (1,2,3), (4,5,6), (7,8,9) are considered as three inputs, the Shor code circuit can be reduced as a sign flip code. In other words, the Shor code can also fix sign flip error for a single qubit.

Any arbitrary error can be corrected by the Shor code, in the case of a single qubit. Considering an error as a unitary transform U (acting on a qubit $|\psi\rangle$), it can be described by the following expression:

$$U = c_0\, I + c_1\, \sigma_x + c_2\, \sigma_y + c_3\, \sigma_z,$$

where c_0, c_1, c_2, c_3 are complex factors, I is the identity operator and there are the familiar Pauli matrices in the formula, namely:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

In case $U = 0$, there was no error. If $U = \sigma_x$, there was a bit flip. In the case of $U = \sigma_z$, there was a sign flip. And if $U = i\sigma_y$, there were both bit flip and sign flip. From the linearity, it follows that the Shor code can fix arbitrary one-qubit errors.

We remark there are other quantum error correcting protocols, but our present knowledge is not enough to understand how they work.

Chapter 10

Quantum search

10.1 The Grover algorithm

The Grover algorithm answers the question how to find a needle in a haystack in a way which is more effective than any of the known classical methods.

What is it all about? Let us supppose we have a list of phone numbers sorted by names and we would like to know whose phone number has been given to us. In the reverse case when we have to find a phone number for a given name, we can use the following *efficient* (classical) algorithm:

We look at the middle of the list of N names and find out if the sought name is before or after the middle of the list. In case it can be found before the middle point, we will look at the middle of the first halfe of the list (at the end point of the quarter of the whole list). Comparing the name found here with the name we seek, we can split the list again to get a new (and smaller) part list to be looked through.

In this way, at the nth step, the length of the part list we have to look through yet is $\frac{N}{2^n}$. That is after approximately $\log_2 N$ steps we find the sought name and the phone number belonging to it. This is an algorithm with a $Poly(\log N)$ complexity which is said to be easy.

However, in the reverse case, it is hard to handle the problem, because the phone numbers are not ordered. If we want to find the name which belongs to the known phone number with a probability of $1/2$, we have to look through $N/2$ names (with their phone numbers), that is we have to take $2^{\log_2(N/2)}$ steps. In other words, in

this case, the task is exponential in $\log N$.

Using the Grover algorithm, the number of steps will be proportional to $\sqrt{N}$ instead of $N/2$. (Though, this effect is not as significant as the accelerator effect of Shor algorithm in the case of prime factorization, but this is an interesting result.)

It is natural to ask how this algorithm works. The mathematical task is the following: Let the list of the names be considered as a set of $Z_N^0 = \{0,\ 1,\ ...\ N-1\}$ and - on this set - let us have the following function $x \in Z_N^0,\ x \to g(x)$, which gives the related phone number $g(x)$. Let us suppose we have been given a phone number $g(\omega)$ and we want to find $\omega \in Z_N^0$, whose image is $g(\omega)$.

We define another function $f(x)$ with the following properties:

$$\begin{aligned} f_\omega(x) &= 1, \quad \text{if} \quad x = \omega \\ f_\omega(x) &= 0, \quad \text{if} \quad x \neq \omega \end{aligned} \tag{10.1}$$

This is the so called oracle, where x is the input and $f(x)$ is the output. Physically, this can be the phone book itself. The essence is the following: for a given x, f can be calculated rapidly, but in the reverse case (where we want to find $x = \omega$ for which $f_\omega(x) = 1$) we face a hard solvable problem.

The real question is that how many times do we need to ask the oracle until we find ω? Let us define the following two-register operation (or "quantum machine"):

$$U_\omega\ :\ |x\rangle|y\rangle \to |x\rangle|y \oplus f_\omega(x)\rangle, \tag{10.2}$$

where, on the first register N different numbers can be imaged, that is the first register is an L-qubit register, where L is the smallest number for which $N < 2^L$ and $|x\rangle$ is imaged as the binary form of number x.

The second register is a two-qubit register. Learning of Deutsch algorithm, we saw that if the initial state of the second qubit is $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, then

$$\begin{aligned} U_\omega\ :\ |x\rangle\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \to \frac{1}{\sqrt{2}}|x\rangle(|0 \oplus f_\omega(x)\rangle - |1 \oplus f_\omega(x)\rangle) = \\ = \begin{cases} \frac{1}{\sqrt{2}}|x\rangle(|0\rangle - |1\rangle) & \text{if} \quad f_\omega(x) = 0 \\ \frac{1}{\sqrt{2}}|x\rangle(|1\rangle - |0\rangle) & \text{if} \quad f_\omega(x) = 1. \end{cases} \end{aligned} \tag{10.3}$$

This means that

$$U_\omega \,:\, |x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \rightarrow (-1)^{f_\omega(x)}|x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{10.4}$$

Now the second register (whose state remains the same) can be omitted and we examine the first one only. Being one of the elements of the special computational base, the sought $|\omega\rangle$ is orthogonal to all the other elements of the base. Thus in case the state of the second register is held unchanged, the effect of operator U_ω can be expressed by the following formula:

$$U_\omega|x\rangle \rightarrow (-1)^{f_\omega(x)}|x\rangle = (1 - 2|\omega\rangle\langle\omega|)|x\rangle, \tag{10.5}$$

because $\langle\omega|x\rangle = 0$, if $\omega \neq x$. Being true for the base vectors, due to the linearity, it has to hold true for the cases of any other vectors. That is operator U_ω has the following shape:

$$\begin{aligned} U_\omega &= 1 - 2|\omega\rangle\langle\omega| \\ U_\omega|\psi\rangle &= |\psi\rangle - 2|\omega\rangle\langle\omega|\psi\rangle. \end{aligned} \tag{10.6}$$

Now we show how the effect of the unitary operator U_ω can be imagined geometrically. In this (geometrical) picture this effect is actually a mirroring to a hyperplane which is perpendicular to $|\omega\rangle$. What does it mean exactly? Let us have the two component vectors of an arbitrary state $|\psi\rangle$. One of them lies in the direction of $|\omega\rangle$ and the other one is perpendicular to this direction. In this case, due to the effect of U_ω the sign of the vector which is parallel with $|\omega\rangle$ will be changed while the other component vector (which is othogonal to $|\omega\rangle$) will be untouched. In case of an arbitrary vector $|\psi\rangle$, a decomposition like this can be seen below:

$$|\psi\rangle = |\omega\rangle\langle\omega|\psi\rangle + (|\psi\rangle - |\omega\rangle\langle\omega|\psi\rangle), \tag{10.7}$$

where the first member is the component which is parallel with $|\omega\rangle$ and the second one is the orthogonal component. Indeed, if the sign of the first member is changed to its reverse (in other words it is mirrored to the plane which is perpendicular to $|\omega\rangle$), the resulted state is

$$|\varphi\rangle = |\psi\rangle - 2|\omega\rangle\langle\omega|\psi\rangle = U_\omega|\psi\rangle, \tag{10.8}$$

which is exactly $U_\omega|\psi\rangle$ (as it can be seen above).

Now let us introduce a state which is the symmetric linear combination of all computational base state:

$$|s\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle. \tag{10.9}$$

We know also $|\omega\rangle$ is one of the elements of the computational base, but we do not know which element is it. Hence, in case of making a measurement in the computational base, the measurement will result $|\omega\rangle$ with a probability of $|\langle\omega|s\rangle|^2 = 1/N$. This probability value is equal to the one which can be obtained in case of seeking $|\omega\rangle$ in a classical random way.

However, Grover algorithm is an iteration, which increases the probability of finding ω step by step. Let us consider the following unitary operator

$$U_s = 2|s\rangle\langle s| - 1. \tag{10.10}$$

Using this operator, its operation preserves the component which lies in the direction of the symmetric average of all base states, that is the component along the direction of $|s\rangle$ will be untuched, while the other component which is orthogonal to $|s\rangle$ will be mirrored.

Now let us consider the Grover operator $G = U_s U_\omega$. Applying it on $|s\rangle$ and again on the resulted state, and so on, after enough number of iterations, we get the proper ω. Indeed, let us consider

$$\langle\omega|s\rangle = \frac{1}{\sqrt{N}} = \sin\theta, \tag{10.11}$$

where the equation defines θ, which is the angle between the symmetric vector and the plane which is orthogonal to the sought $|\omega\rangle$. Let us have two orthogonal components of $|s\rangle$, where one of them is in the direction of $|\omega\rangle$ (and naturally the other one is orthogonal to it). In this way we get the following expression:

$$|s\rangle = |\omega\rangle \sin\theta + |\omega_\perp\rangle \cos\theta, \tag{10.12}$$

where $|\omega_\perp\rangle$ is a unit vector which lies in the direction of $|s_\perp\rangle = |s\rangle - |\omega\rangle\langle\omega|s\rangle$, that is $|\omega_\perp\rangle = |s_\perp\rangle/|s_\perp|$. Applying U_ω on $|s_\perp\rangle$, we get the following result: $|s'\rangle = -|\omega\rangle \sin\theta + |\omega_\perp\rangle \cos\theta$. If we apply U_s on the previous result, overall we achieve the following operation: $G|s\rangle$, or

$$G|s\rangle = |s_1\rangle = |\omega\rangle \sin 3\,\theta + |\omega_\perp\rangle \cos 3\,\theta. \tag{10.13}$$

(This result can be seen geometrically, but also can be deducted from the formulae above.) From this, it directly follows that vector $|s\rangle$ turns towards $|\omega\rangle$ from $|\omega_\perp\rangle$. It can be seen easily that if we apply G again, the resulted state will be

$$G^2|s\rangle = G|s_1\rangle = |s_2\rangle = |\omega\rangle \sin 5\,\theta + |\omega_\perp\rangle \cos 5\,\theta \tag{10.14}$$

and in general

$$G^n|s\rangle = |s_n\rangle = |\omega\rangle \sin(2n+1)\,\theta + |\omega_\perp\rangle \cos(2n+1)\,\theta. \tag{10.15}$$

If we stop iterating in the proper moment, then $|s_n\rangle$ will be close to $|\omega\rangle$, that is $\langle s_n|\omega\rangle \approx 1$. This means if we make a measurement in the computational base we will get $|\omega\rangle$ with a big probability. This can happen, if the angle of the turning is about $\pi/2$.

If N is big, then $\sin\theta = 1/\sqrt{N} \ll 1$ so $\theta \approx \sin\theta = 1/\sqrt{N}$. Thus at the Tth step the angle is $(2T+1)\,\theta \approx (2T+1)/\sqrt{N}$ and this value is close to $\pi/2$, that is $(2T+1)/\sqrt{N} \approx \pi/2$. From this it follows that the number of the steps we need to take is about the following value:

$$T = \frac{\pi}{4}\sqrt{N} - \frac{1}{2}, \tag{10.16}$$

which is proportional to $\sqrt{N}$. However, T has to be an integer number so the following formula is righter: $T = \frac{\pi}{4}\sqrt{N}(1 - O(1/\sqrt{N}))$. After taking these steps the probability of finding ω is

$$P(\omega) = |\langle s_n|\omega\rangle|^2 = \sin^2(2T+1)\,\theta = 1 - O(1/N). \tag{10.17}$$

Operator U_ω has to be applied once in every step, hence also function f has to be examined once in every step. The previous result means that our search is $\sqrt{N}$ times shorter than in a classical case. However, we have to be careful, because in case we do not stop iterating in time, we will move away from ω.

Let us consider a simple case, where $N = 4$. Though, in this situation θ is not small, but our method works well, because $\sin\theta = 1/2$ and from this it directly follows $\theta = \pi/6$, that is after one rotation the value of the angle $3\,\theta = \pi/2$. This means that the result is equal to ω exactly. If we wanted to solve this problem using classical search, the expectation value of the number of tryings would be $1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} = 2.5$

Chapter 11

Universal covariant quantum cloner

Though - as we already know - quantum cloning is an impossible quantum map, it is still possible to create imperfect replicas of a quantum state with optimal fidelity [4]. This protocol, originally introduced by Bužek and Hillery [5], is called quantum cloning. It is also possible to perform asymmetric quantum cloning: the fidelity of the copies might not be equal [6, 7]. Quantum cloning has been studied very extensively in the literature, and it has many variants. See Ref. [8] for a review.

In this chapter we are interested in asymmetric universal cloning transformations for individual quantum bits. A quantum circuit was designed by Bužek et al. for this purpose [9], which was later generalized to arbitrary dimensional quantum systems [10]. We shall call this circuit UCQC (universal covariant quantum cloner) in what follows. It has a special feature of being quantum controlled, that is, the fidelity ratio of the two clones is controlled by the initial quantum state of two ancillary quantum bits (one of which will carry the clone after the process). This idea turned out to be related to the concept of programmable quantum networks or quantum processors [11]. These are fixed quantum networks which are capable of performing operations on quantum systems in a way that the operation itself is encoded into the initial quantum state of ancillae. It was found that the very circuit for universal quantum cloning is in fact a probabilistic universal quantum processor [12].

In this chapter we consider UCQC-s as entanglement manipulation devices. In the context of cloning, one may ask several questions. One may consider the cloning of an entangled quantum state as a whole, in order to obtain similar entangled pairs. For two qubits this has been analyzed in detail by several authors [13, 14,

15]. In particular, Bužek at al. [16] compare the fidelity of cloning of an entangled pair by global and local operations.

Another approach might be the broadcasting of entanglement, proposed by Bužek et al. [16]. In this case two parties share an entangled pair and use cloners locally to obtain two partially entangled pairs. This protocol attracted a relevant attention in the literature, too. Topics such as state-dependent broadcasting [17], broadcasting of multipartite entangled states: W-states [17, 18], GHZ-states [19], and linear optical realizations [20] were discussed in detail.

In this chapter, we consider the entanglement manipulation capabilities of the universal covariant quantum cloner or quantum processor circuit for quantum bits. We investigate its use for cloning a member of a bipartite or a genuine tripartite entangled state of quantum bits. We find that for bipartite pure entangled states a nontrivial behavior of concurrence appears, while for GHZ entangled states a possibility of the partial extraction of bipartite entanglement can be achieved.

Consider an entangled pair. It is always interesting to ask what happens to the entanglement if any of the members of the pair is subjected to some quantum information processing protocol. In the case of quantum teleportation, for instance, rather strikingly the teleported qubit inherits the entanglement of the original qubit with its pair. It is rather natural to ask what happens in the case of a universal quantum cloner. The answer for qubit pairs is partly given by Bandyopadhyay and Kar [21]. They show that if a member (or both members) of a maximally entangled qubit pair is subjected to an optimal universal quantum cloning operation, the resulting state is a Werner state. It is likely, however, that a cloning transformation is realized by some quantum circuit, which uses ancillae for carrying out the operation. It is obviously interesting how the entanglement between the different quantum bits of such a scenario (including also ancillae) behaves. In this chapter we consider the UCQC as a circuit, not only the cloning operation itself. We calculate entanglement as measured by concurrence (see Chapter 8). It turns out that the ancillae play a very specific role and the behavior of concurrence shows a rather interesting pattern. The recent optical realization of certain programmable quantum gate arrays [22] also contributes to the relevance of this question.

Another similar question might be the partial extraction of bipartite entanglement from a GHZ-type threepartite resource. It is known that if three qubits are in a GHZ state [23], then a measurement on either of the three qubits in the $|\pm\rangle$ basis (eigenbasis of the σ_x Pauli-operator) projects the state of the remaining two qubits into a maximally entangled state. We show that if the given particle is cloned in advance, it is possible to create bipartite entanglement by measuring the clone, while there still remains some purely threepartite entangled resource in the state

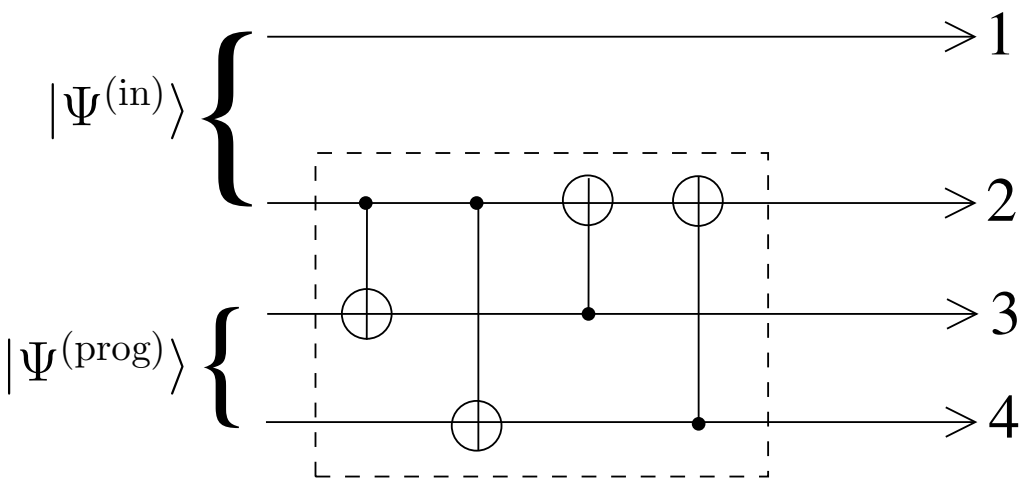


Figure 11.1: The setup for bipartite entangled states. The dashed box contains the universal covariant cloning circuit, composed of four controlled-NOT gates.

of the three parties. This is indicated by the possibility of entangling a different pair of qubits by a next measurement. The nature of the entanglement in the multipartite system can be also analyzed with the aid of the Coffman-Kundu-Wootters inequalities [2] (see Chapter 8), which quantify the monogamy of entanglement. We shall present such an analysis, too.

This chapter is organized as follows: in Section 11.1 we analyze the behavior of bipartite entanglement in the case when UCQC is applied to clone a member of a maximally entangled pair. In Section 11.2 we consider the application of UCQC for the partial extraction of bipartite entanglement from a Greenberger-Horne-Zeilinger state. In Section 11.3 the results are summarized and conclusions are drawn.

11.1 Bipartite pure states

The considered setup is depicted in Fig. 11.1. The quantum circuit in the dashed box is the universal quantum cloner [10]. Its first input port (which is port 2 in our current notation) receives the state to be cloned, while on the second two ports there impinges a so-called program state:

$$|\Psi_{34}^{(\mathrm{prog})}\rangle = \mathcal{N}\left(\alpha\left(|0\rangle(|0\rangle+|1\rangle)\right)+\beta\left(|00\rangle+|11\rangle\right)\right), \tag{11.1}$$

where $\mathcal{N} = 1/\sqrt{2(\alpha+\beta^2)}$ is a normalization constant, α is a real parameter and $\beta = 1-\alpha$. (In Ref. [9] there is an additional circuit introduced to prepare the program state of Eq. (11.1), as a first step of cloning, which we have omitted here, as it does not have a role in the cloning process itself.) Were a single-qubit state ϱ impingent on port 2, the output states would be:

$$\begin{aligned}
\varrho_2 &= \frac{\beta}{\alpha+\beta^2}\varrho + \frac{\alpha^2}{2(\alpha+\beta^2)}\hat{1}, \\
\varrho_3 &= \frac{\alpha}{\beta+\alpha^2}\varrho + \frac{\beta^2}{2(\beta+\alpha^2)}\hat{1}, \\
\varrho_4 &= \frac{\alpha\beta}{\beta^2+\alpha}\varrho^T + \frac{\alpha^2+\beta^2}{2(\alpha+\beta^2)}\hat{1}.
\end{aligned} \tag{11.2}$$

The clones reside in ports 2 and 3, the original qubit and the first ancilla, whereas in the port 4 there is an ancilla, the state of which is a mixture of the state described by the mixture of the transpose of the density operator of the original state and the identity operator. The fidelity of the clones depends on the value of α: for $\alpha = 0$ there is no cloning, whereas for $\alpha = 1$ the state of the original qubit is fully transferred to the clone, leaving the original qubit in a completely mixed state. For other values of alpha there are optimal clones generated. Note the symmetry of the formulae in α and β.

In the terminology of cloning this setup realizes an optimal universal asymmetric cloner. The term asymmetric expresses that the two clones are not identical, their fidelity with respect to the original state is different, but controlled by the parameter α. Setting $\alpha = \frac{1}{2}$, we obtain the symmetric case. The cloning is universal in the sense that the realized cloning transformation itself does not depend on the input state. Optimality in this context means that the second clone is obtained with maximal fidelity for a given fidelity of the first one.

Let us return to the description of the whole scenario in argument, depicted in Fig. 11.1. The qubits 1 and 2 carry the initial bipartite input state. Qubit 2 is subject to cloning, while qubit 1, the first part of the pair is not manipulated. We are interested in the entanglement relations between the different pairs of qubits in the resulting state. As for the measure of bipartite entanglement for qubits, we apply concurrence according to the Wootters formula as we saw in Eq. (8.17).

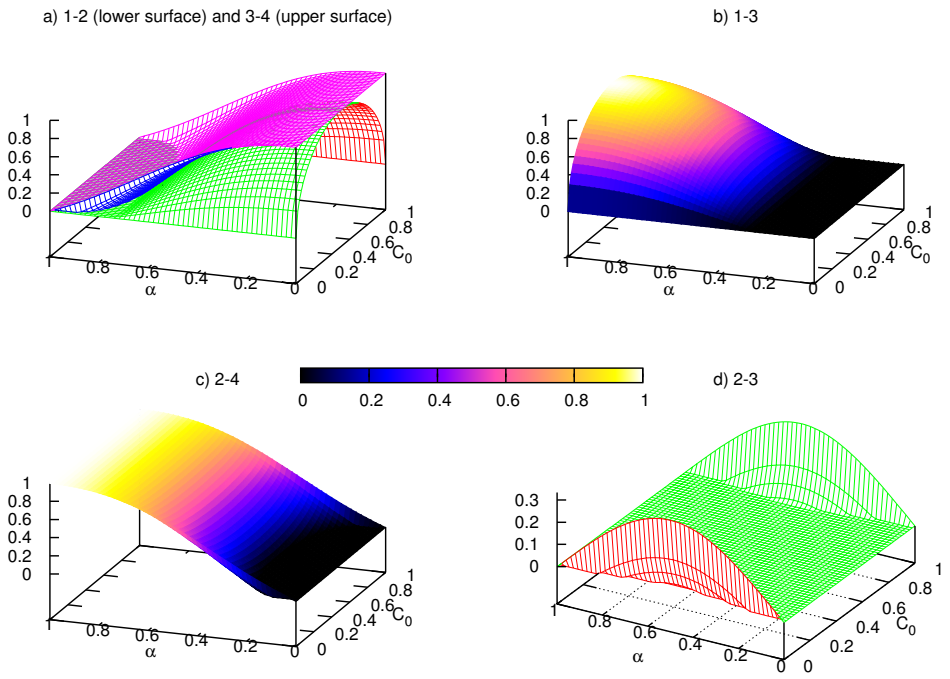


Figure 11.2: (color online) The concurrence between the various pairs of qubits at the output of the setup in Fig. 11.1. The input state is either of the four in Eq. (11.3), the same figure is obtained for each choice, though the states themselves differ. The "program" state is the one in Eq.(11.1). The concurrence between qubits 1 and 4 is zero. The plotted quantities are dimensionless.

As an input state we consider a state in either of the following four forms:

$$|\Phi_{12}^{(\mathrm{in})}\rangle = \sqrt{C_0}|00\rangle \pm \sqrt{C_1}|11\rangle,$$
$$\text{or}$$
$$|\Psi_{12}^{(\mathrm{in})}\rangle = \sqrt{C_0}|01\rangle \pm \sqrt{C_1}|10\rangle,$$
$$C_0 + C_1 = 1. \tag{11.3}$$

As for the nonzero concurrences between the various pairs of qubits, we obtain the behavior in Fig. 11.2, regardless of the choice from the above states. The output states, however, depend on this actual choice, we shall comment on this later.

In the figure one can observe that the entanglement between qubits 1 (the one not manipulated) and qubit 2 (the original qubit) behaves in the similar way as that between qubit 1 and 3 (the one not manipulated and the clone). For $\alpha = 0$ (no cloning), qubits 1 and 2 are entangled as they were originally, while for $\alpha = 1$,

complete cloning, the entanglement is transferred to qubits 1 and 3, the clone plays the role of the former original qubit completely. The surfaces representing the concurrence for the pairs 1-2 and 1-3 are symmetric in the cloning parameter α, that is, they can be obtained from each other by the $\alpha \to 1-\alpha$ substitution. The dependence of these entanglements from α is monotonous but not continuous: for small values there is a region where the entanglement is zero, and it appears suddenly and non continuously. The dependence of these concurrences on the initial entanglement in the state C_0 is monotonous and continuous.

It is also interesting to observe that a similar non-symmetric behavior appears in the concurrence of qubits 3-4 and 2-4. The program state of Eq. (11.1), in which the qubits 3-4 are prepared initially, is maximally entangled for $\alpha = 0$, the case of no cloning, and its entanglement decreases with the increase of the cloning parameter α. Accordingly, the entanglement of qubits 3-4 decreases with α also after the cloning operation, while the complementary behavior (in the sense of $\alpha \to 1-\alpha$ substitution) appears between qubits 2 and 4 (the cloned part of the input state and the ancilla of the cloner). Note that the entanglement of qubits 3 and 4 is not equal to their entanglement *before* the cloning operation: the concurrence of the partially entangled program state in Eq. (11.1) is a monotonous and continuous function of α, and its values are not equal to the concurrences after the cloning operation. Moreover, the concurrence of qubits 3 and 4 after the cloning also depends slightly on that of the input state of qubits 1 and 2, in Eq. (11.3).

As for the remaining pairs, qubits 1 and 4 (the qubit not manipulated and the ancilla) will not be entangled, while between qubits 2 and 3 (the second input state and its clone), as a nontrivial effect, there is a small amount of entanglement appearing only in the case the input state (of qubits 1-2) is only slightly entangled.

A special case arises if the input state of qubits 1 and 2 is maximally entangled. This is the case of $C_0 = 1/2$ in Fig 11.2. The concurrence between qubits 1-2 and 3-4 (two originals, two program qubits of the cloner) is equal to each other. The complementary pairs, qubits 1-3 (not manipulated-clone) and 2-4 (clone-ancilla) have also equal concurrences. The dependence of these concurrences on α is depicted in Fig. 11.3. The behavior of these curves is due to the fact that the universal cloning transformation produces Werner states. Indeed, if the input state is in Eq. (11.3) is the maximally entangled $|\Phi^{(+)}\rangle$ Bell-state, where the states of qubit-pairs 1-2, 1-3, 2-4, 3-4 are Werner-states of the form

$$\varrho^{(\text{Werner})} = \gamma |\Phi^{(+)}\rangle\langle\Phi^{(+)}| + \frac{1-\gamma}{4}\hat{1}, \tag{11.4}$$

where $\hat{1}$ stands for the identity operator of the two-qubit space. The value of the

parameter γ is

$$\gamma_{12} = \gamma_{34} = \frac{\alpha}{\alpha + \beta^2} \tag{11.5}$$

for pairs 1-2 and 3-4, while it is

$$\gamma_{13} = \gamma_{24} = \frac{\beta}{\alpha + \beta^2}. \tag{11.6}$$

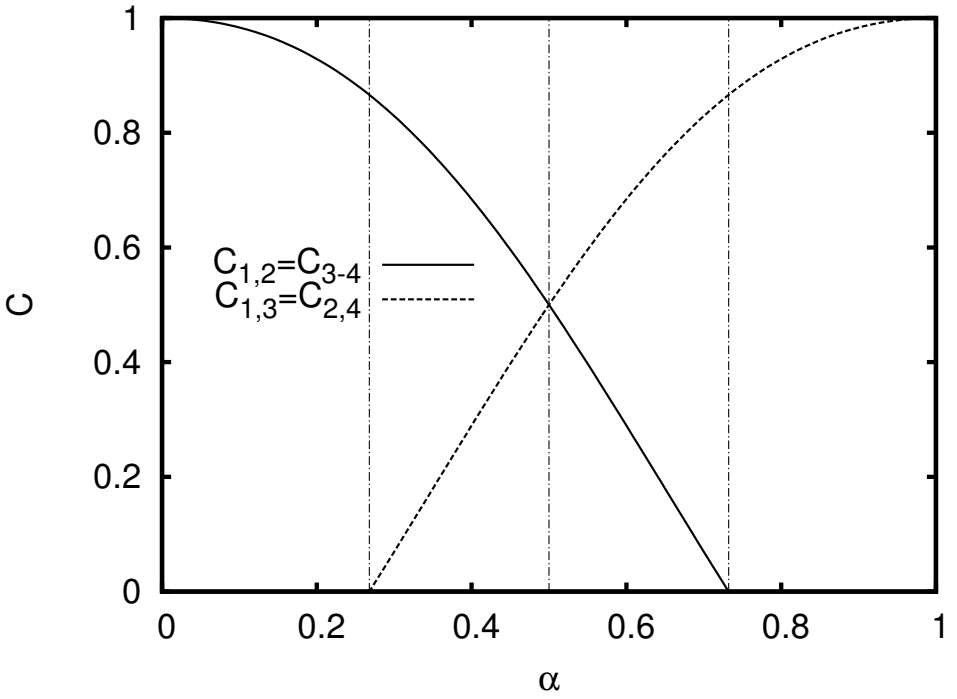


Figure 11.3: A slice of Figs. 11.2 a-c) for $C_0 = 1/2$, that is, for any of the maximally entangled Bell-states as input. The plotted quantities are dimensionless.

Note that the denominator on the right-hand-side of the above formulae comes directly from the normalization constant of the program state in Eq. (11.1) (i.e. the scaling of parameters in Eqs. (11.5) and (11.6) is merely a consequence of our particular choice of parameters). In the case we choose a different one from the states in Eq. (11.3), we obtain local unitary transforms of the Werner state in Eq. (11.5). The message of the consideration for cloning an element of a maximally entangled pair is not the fact that Werner states are obtained in qubits 1-2 and 1-3, since it was known from the literature [21]. What is nontrivial here that in the UCQC circuit this behavior is repeated between the ancilla (qubit 4), and qubits 2 and 3, and this holds only in the case of the cloning of a member of a maximally entangled state. Finally let us note that the behavior of qubits 2, 3, and 4 cannot depend on the properties of qubit 1 since it is a remote system from the UCQC's point of view. It is the reduced density operator of qubit 2 which can influence their behavior. We have found that only for a maximally entangled pair, a concurrence characterizing a nonlocal property is equal to another concurrence which is a local property of the cloner.

11.2 The GHZ state

In this section we consider the case in which a member of a Greenberger-Horne-Zeilinger (GHZ) state is cloned. This tripartite state, of the form

$$|\Psi^{(GHZ)}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \tag{11.7}$$

is known to be genuinely tripartite entangled. That is, all the pairwise entanglements (as measured by concurrence) are zero, however, all of the three qubits are in a maximally entangled state. When any of the qubits is subject to a von Neumann measurement in the basis

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle), \tag{11.8}$$

the other two qubits will be in either of the maximally entangled Bell-states

$$|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \tag{11.9}$$

depending on the measurement result. The probability of the measurement results are equal. In this way the tripartite entangled resource in the GHZ state can be converted into maximal bipartite entanglement.

The scenario we consider for GHZ states is depicted in Fig. 11.4. Qubits 1-3 carry the input state which is a GHZ state in Eq. (11.7). Qubit 3 enters the UCQC's first port. The program ports of the UCQC are qubits 4 and 5, considered again to be in the program state in Eq. (11.1).

Directly after the operation of the cloner all pairwise concurrences are zero, except for the one between qubits 3-5 and 4-5. Their value is represented by curves "A" and "B" in Fig. 11.5. This fact is easily explained by the following reasoning. From the point of view of the UCQC circuit, qubits 1 and 2 are remote ones, thus they cannot influence the local properties of qubits 3, 4 and 5. All we "see" at the locus of the UCQC is that qubit 3 is in a maximally mixed state, as it is a member of the maximally entangled threepartite GHZ state of Eq (11.7). But the same situation would arise if qubit 3 were maximally entangled in a bipartite sense with one additional qubit, as we have considered in the previous Section. Thus the behavior of concurrences between the pairs ancilla-original and ancilla-clone are the very same as in the case of cloning a member of a bipartite maximally entangled state (or a member of any kind of multipartite entangled state which is itself in a completely mixed state for this reason): Werner states are obtained.

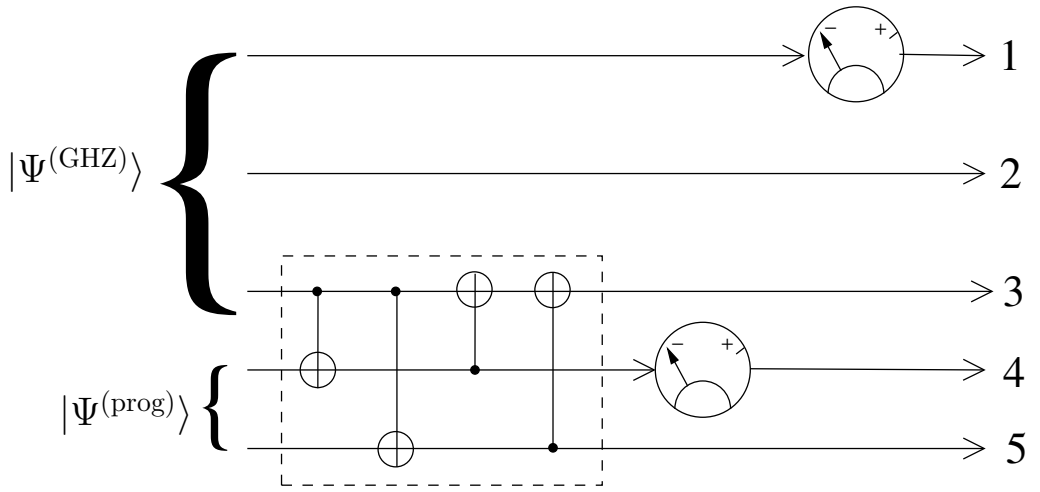


Figure 11.4: The setup for the tripartite GHZ state. Qubits 1-3 hold the GHZ state initially. The dashed box contains the universal covariant cloning circuit, composed of four controlled-NOT gates. The "meters" measure in the $|\pm\rangle$ basis. First the clone (qubit 4), then a member of the original GHZ state (qubit 1) is measured. The horizontal position from the left to the right side thus represents the order of operations in time.

Projective measurement on the clone. Motivated by the relation of the projective measurements of the members of a GHZ state on the $|\pm\rangle$ basis, one may now consider a measurement of this kind on the clone, that is, on qubit 4. This measurement will not alter the bipartite entanglement between qubits 3-5, and that between 4-5 will disappear due to the measurement. However, there will be an even larger entanglement appearing between qubits 1 and 2, this is curve "C" in 11.5. Both measurement outcomes will have equal probability and also the entanglement behavior is the same for both cases. In case of full cloning ($\alpha = 1$), we obtain a pure EPR pair as expected. (We remark here that if we were to measure on the original qubit (qubit 3) instead of its clone, we would obtain the counter propagating curve of the same shape, curve "D" in Fig. 11.5, as one would expect. The role of the original and the clone is symmetric. Entanglement of 4-5 will not alter, while that of 3-5 will disappear in this case.) This is a partial conversion of the resource available as genuine tripartite entanglement into bipartite entanglement.

Measurement on the qubit 1. In order to further justify this statement let us consider a second measurement, now on qubit 1. Again, the results will be uniformly distributed and the entanglement itself will not depend on the measurement

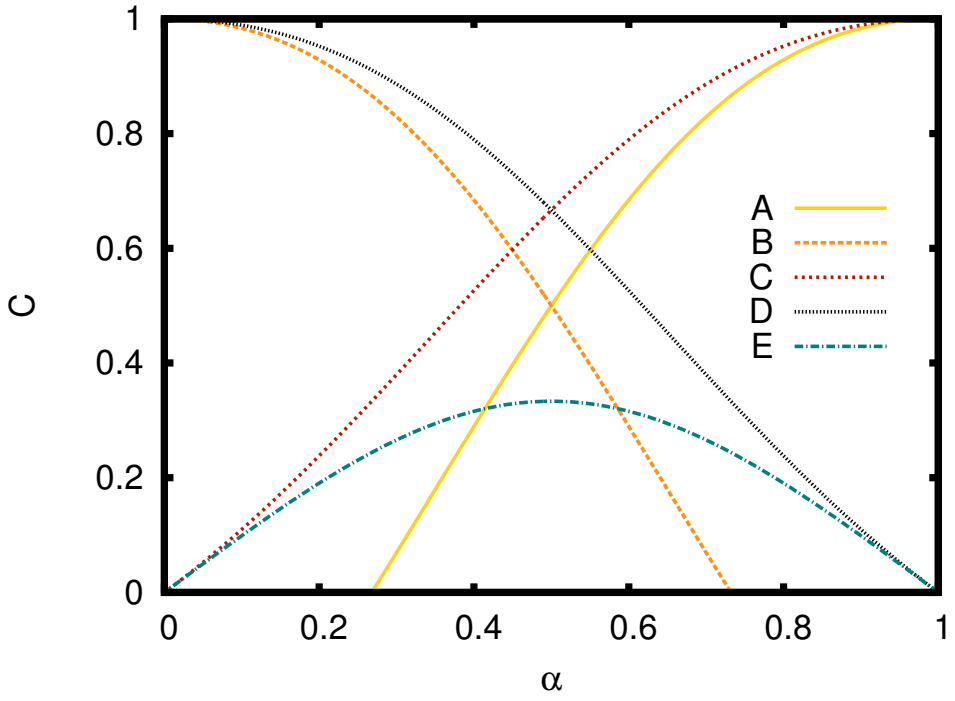


Figure 11.5: (color online) Pairwise concurrences in the GHZ-cloning scenario. A: qubits 3-5 after the cloner and also after each measurement, B: qubits 4-5 after the operation of the cloner, C: qubits 1-2 after the first measurement, D: qubits 2-3 after the second measurement, E: qubits 2-5 after the second measurement. The plotted quantities are dimensionless.

result. The entanglement between qubits 3 and 5 will be untouched, and that between qubits 1 and 2 will be destroyed by the measurement of course. Meanwhile we obtain nonzero entanglement between pairs 2-3 and 2-5, these are the curves "D" and "E" in Fig. 11.5, respectively. Indeed, if the extraction of the tripartite entanglement was not full (i.e. $\alpha \neq 1$), one can still obtain bipartite entanglement by measuring another qubit this time. Curves "C" and "D", describing the entanglement between 1-2 after the first measurement, and that between 2-3 after the second measurement, respectively, are counter propagating, reflecting the interplay between the two extractions. As a side effect, there is a small amount of entanglement which appears between qubits 2-5 after the second measurement, this is curve "E" in Fig. 11.5.

The use of the partial extraction of the entanglement is the following. Consider that qubit 1 is at Alice, qubit 2 at Bob, while the rest of the qubits is at Charlie. Initially they share a tripartite GHZ resource. Charlie wants to enable Alice and Bob to use a bipartite maximally entangled channel. He might perform the projective measurement on the clone he has, however, in this case his qubit 3 gets disentangled from the rest of the parties. However, if he performs cloning and measures the clone, Alice and Bob still obtains a partially entangled bipartite resource. However, Alice can decide that instead of using a bipartite channel with Bob, she wants to create a channel between Bob and Charlie. All she has to do is to perform a proper measurement on her qubit and communicate the result: Bob and Charlie shall posses a partially entangled bipartite resource. This would not be possible without the cloning. The same could be done of course by Bob, to enable the bipartite resource between Alice and Charlie.

In order to obtain a deeper insight into the behavior of bipartite entanglement in this multipartite system, it is worth examining the Coffman-Kundu-Wootters inequalities. As we mentioned in Chapter 8, if inequalities in Eq. (8.20) are saturated, the bipartite entanglement is maximal.

To quantify the saturation we evaluate

$$s = \tau_k - \sum_{l \neq k} C_{k,l}^2, \tag{11.10}$$

which is zero if the inequalities are saturated. After the first measurement we obtain nonzero values except for the fourth qubit (apart from the case of $\alpha = 1$. The behavior is depicted in Fig. 11.6. The fact that the CKW inequalities are not saturated also suggests the presence of additional multipartite entanglement in the system. After the second measurement, on the other hand, we find that all the CKW inequalities are saturated: the system is in a sense maximally bipartite entangled.

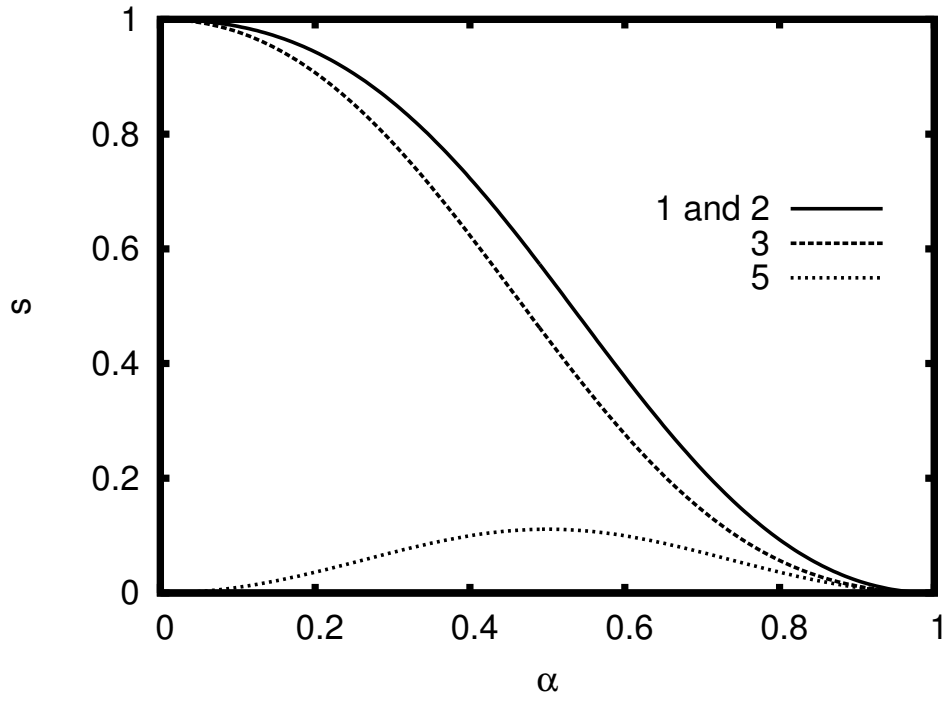


Figure 11.6: The quantity in Eq. (11.10), which is zero if the CKW inequalities are saturated, for each qubit after the first measurement in our GHZ-cloning scenario. For qubit 4 the quantity is zero. The plotted quantities are dimensionless.

11.3 Conclusion

We have shown that when using a universal covariant quantum cloning circuit to clone a member of an entangled pair of qubits, a very specific behavior of the entanglement of the qubits appears. The main feature is that behavior of the entanglement between the not cloned part of the pair and the cloned one is repeated in the entanglement of certain ancillae, and so is that of the not cloned qubit and the clone, provided that the original qubit pair was maximally entangled initially. We have described the behavior of the entanglement in detail.

We have also investigated the cloning of an element of the GHZ state. It appears that the universal quantum cloning circuit facilitates the partial extraction of bipartite entangled resources from a genuine tripartite entangled resource. We provided a detailed analysis of the entanglement behavior, including the relation to Coffman-Kundu-Wootters inequalities.

In conclusion, the universal quantum cloning circuit (or quantum processor) for qubits is found to be useful as an entanglement manipulator as well. It can perform entanglement manipulations which are potentially applicable in quantum information processing.

Bibliography

[1] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[2] V. Coffman, J. Kundu, and W. K. Wootters, Phys. Rev. **61**, 052306 (2000).

[3] W. K. Wooters and W. H. Zurek, A single quantum cannot be cloned, Nature **299**, 802 (1982)

[4] R. F. Werner, Phys. Rev. A **58**, 1827 (1998).

[5] V. Bužek and M. Hillery, Phys. Rev. A **54**, 1844 (1996).

[6] N. J. Cerf, Phys. Rev. Lett. **84**, 4497 (2000).

[7] N. J. Cerf, J. Mod. Opt. **47**, 187.209 (2000).

[8] V. Scarani, S. Iblisdir, N. Gisin, and A. Acín, Rev. Mod. Phys. **77**, 1225 (2005).

[9] V. Bužek, S. L. Braunstein, M. Hillery, and D. Bruß, Phys. Rev. A **56**, 3446 (1997).

[10] S. Braunstein, V. Bužek, and M. Hillery, Phys. Rev. A **63**, 052313 (2001).

[11] M. A. Nielsen and I. L. Chuang, Phys. Rev. Lett. **79**, 321 (1997).

[12] M. Hillery, V. Bužek, and M. Ziman, Phys. Rev. A **65**, 022301 (2002).

[13] J. Novotný, G. Alber, and I. Jex, Phys. Rev. A **71**, 042332 (2005).

[14] L.-P. Lamoureux, P. Navez, J. Fiurášek, and N. J. Cerf, Phys. Rev. A **69**, 040301 (2004).

[15] E. Karpov, P. Navez, and N. J. Cerf, Phys. Rev. A **72**, 042314 (2005).

[16] V. Bužek and M. Hillery, Phys. Rev. Lett. **81**, 5003 (1998).

[17] S. Adhikari, B. Choudhury, and I. Chakrabarty, J. Phys. A-Math. Gen. **39**, 8439 (2006).

[18] I. Chakrabarty and B. S. Choudhury, Int. J. Quantum Inf. **7**, 559 (2009).

[19] S. Adhikari and B. S. Choudhury, Phys. Rev. A **74**, 032323 (2006).

[20] I. Ghiu and A. Karlsson, Phys. Rev. A **72**, 032331 (2005).

[21] S. Bandyopadhyay and G. Kar, Phys. Rev. A **60**, 3296 (1999).

[22] M. MiCuda, M. Ježek, M. Dušek, and J. Fiurašek, Physical Review A (Atomic, Molecular, and Optical Physics) **78**, 062311 (2008).

[23] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Physics Today **46**, 22 (1993).

[24] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[25] J. S. Bell, Physics (NY) **1**, 195 (1965).

[26] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[27] L. Hardy, Phys. Rev. Lett. **68**, 2981 (1992).

[28] J. S. Lundeen and A. M. Steinberg, Phys. Rev. Lett. **102**, 020404 (2009).

[29] K. Yokota, T. Yamamoto, and M. K. N. Imoto, New J. Phys. **11**, 033011 (2009).

[30] T. Geszti, Phys. Rev. A **58**, 4206 (1998).

[31] M. Koniorczyk and V. Bužek, Phys. Rev. A **71**, 032331 (2005).

[32] T. Laustsen, F. Verstraete, and S. J. van Enk, Quantum Inf. Comput. **3**, 64 (2003).

[33] M. Koniorczyk, L. Szabó and P. Adam, Phys. Rev. A **84**, 044102 (2011).

[34] A. C. Elitzur and L. Vaidman, Found. Phys. **23**, 987 (1993): L. Vaidman, Quantum Opt. **6**, 119 (1994).