Experiments that gave rise to the development of Quantum Mechanics and experiments that have further underpinned the theory

by

Jens Hee
https://jenshee.dk

October 2023

Change log

31. October 2023

1. Document started.

Contents

Chapter 1

Introduction

Quantum Mechanics is the theory about the behaviour of particles at the atomic and subatomic level. In fact, phenomena at the atomic scale can only be explained by Quantum Mechanics. However, Quantum Mechanics is applicable to macroscopic phenomena as well, though it is not commonly used due to the extremely small quantum effects in such cases. A consequence of physical experiments carried out in the 20th century and various theories based on these experiments, led to the works in 1926 by Werner Heisenberg and Erwin Schrödinger and Quantum Mechanics was born.

More experiments have been made up till today, and they have underpinned the correctness of Quantum Mechanics. Some of the results of these experiments are so counter-intuitive that several physicists have claimed that Quantum Mechanics must be incomplete, but no experiment has been made that disproves the theory from 1926. There have also been several attempts to give an interpretation of the theory and the strange results found by experiments. Among these interpretations, the so-called Copenhagen interpretation (see appendix A) is widely accepted. In the following some of these experiments are described along with some theoretical explanation.

It is the author's experience that it is important to study the experiments before going into details with the theory of Quantum Mechanics. More details on the experiments can be found in textbooks and on the net; for example, Wikipedia has several in-depth articles on the subjects. References to these articles are given at the end of this paper.

To answer your question, quantum mechanics and classical physics are two mathematical models of reality that describe different aspects of nature.

Quantum mechanics applies to microscopic systems, such as atoms and subatomic particles, while classical physics applies to macroscopic systems, such as planets and stars. Quantum mechanics and classical physics have different mathematical formalisms and make different predictions about the behavior of physical systems.

For example, quantum mechanics predicts that energy, momentum, and other quantities are quantized, meaning that they can only take discrete values, while classical physics assumes that they are continuous.

Quantum mechanics also predicts that physical systems can exhibit both particle-like and wave-like properties, depending on how they are measured, while classical physics treats them as distinct categories.

Furthermore, quantum mechanics introduces the uncertainty principle, which states that there are limits to how precisely one can know the values of certain physical quantities, such as position and momentum, at the same time, while classical physics assumes that they can be measured with arbitrary accuracy. Therefore, quantum mechanics and classical physics are two mathematical models of reality that are valid in different domains and have different assumptions and implications. They are not contradictory, but rather complementary, as most classical physics theories can be derived from quantum mechanics as approximations in the limit of large scales.

Chapter 2

The Black Body Radiation

A black body is an object that absorbs all the incoming electromagnetic waves. It is only black at 0 °K. At higher temperatures it emits electromagnetic energy with a maximum energy density depending on the temperature and the wavelength, see Figure 2.1 The phenomenon is well known from heated iron, it turns red when the temperature reaches some hundred ℃. Experiments have shown that the color is independent of the material the black body is made of. In the nineteenth century the phenomenon was studied by several physicists, but no theory could explain the phenomenon. On the contrary theories led to an increasing energy density with frequency, which is in clear contrast with experiments and common sense. This is called the ultraviolet catastrophe.
Around 1890 the physicist Max Planck became intrigued by the problem and decided to solve it. He first made some attempts to find an empirical formula that would fit the measured data. After more accurate data became available especially at low frequencies, he succeeded in finding a satisfactory formula.

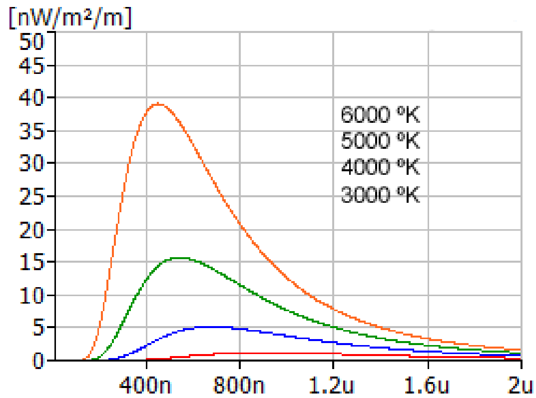

Figure 2.1: The Black Body radiation

$$I_\lambda(\lambda) = \frac{2hc^2}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1} \tag{2.1}$$

But Max Planck was not satisfied with a formula that was not founded in a physical theory an used several years before he in 1901 published a paper where he presented a theory that explained the phenomenon. What he did was to work backwards from his empirical formula and managed, using statistical mechanics, to find a strange explanation:

Electromagnetic waves always have energies equal to hν, where h is a constant now called the Planck constant and ν is the frequency of the electromagnetic wave.

As we say today the energy is quantized, but it is a continuous function of frequency. Max Plank did not at that time regard the theory as final, he could not imagine that electromagnetic waves in reality had this behaviour. But he was convinced that he was on to something important.
As explained in the next sections it was indeed a breakthrough in modern physics and was the first step in the development of Quantum Mechanics. The next section utilizes statistical mechanics to derive the energy density function of a black body, incorporating our contemporary understanding of quantum mechanics.

Links:
https://en.wikipedia.org/wiki/Black-body_radiation
https://en.wikipedia.org/wiki/Max_Planck

Black Body

In a cavity the energy density $(J/m^3/Hz)$ is independent of the shape of the cavity, since two cavities connected by a thin channel exchange a zero net energy flow if the temperature is the same.
Given a box cavity (L_x, L_y, L_z) at temperature T then the electromagnetic modes are given by:

$$A\sin(\frac{n_x\pi x}{L_x})\sin(\frac{n_y\pi y}{L_y})\sin(\frac{n_z\pi z}{L_z})\sin(2\pi\nu t)$$

The composite wave number is given by:

$$(k_x, k_y, k_z) = (\frac{n_x}{2L_x}, \frac{n_y}{2L_y}, \frac{n_z}{2L_z})$$

and the magnitude is:

$$k = \sqrt{k_x^2 + k_y^2 + k_z^2}$$

k defines a regular lattice where each box has the volume $\frac{1}{2L_x 2L_y 2L_z} = \frac{1}{8V}$ and the number of modes within $[k; k+dk]$ is $\frac{1}{8}$ of the surface area of a sphere with radius k, divided by the box volume times dk, since $k_x, k_y, k_z > 0$.

$$dN(k) = \frac{\frac{1}{8}4\pi k^2}{\frac{1}{8V}}dk = 4\pi V k^2 dk$$

Since $\nu = kc$, where c is the speed of light, the number of modes per unit volume within $d\nu$ is given by:

$$dN(\nu) = \frac{8\pi}{c^3}\nu^2 d\nu$$

since the radiation is polarized in two direction.

Energy density inside a cavity

The energy density inside a cavity is found by multiplying the number of modes per unit volume with the energy of the harmonic oscillator (see the next section):

Classical case

$$u_\nu(\nu) = \frac{8\pi}{c^3}\nu^2 k_B T = \frac{8\pi}{c^3}\nu^2 k_B T$$

The unit is $J/m^3/Hz$. k_B is the Boltzmann constant and T is the temperature in degrees Kelvin. It is seen that the energy density is increasing with frequency, i.e the ultra violet catastrophe.

Quantum case

$$u_\nu(\nu) = \frac{8\pi}{c^3}\nu^2 \frac{h\nu}{e^{\frac{h\nu}{k_B T}} - 1} = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{\frac{h\nu}{k_B T}} - 1}$$

The unit is $J/m^3/Hz$.
Often the energy density is given as a function of wavelength:

$$u_\lambda(\lambda) = u_\nu(\frac{c}{\lambda})\frac{\frac{c}{\lambda}}{\lambda} = \frac{8\pi hc}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1}$$

The unit is $J/m^3/m$.

Intensity of the radiation

In this section the cavity is regarded as the enclosure of a photon gas, where the photons do not interact with each other and the trajectories are straight lines. Let the radiation intensity from the inner surface of the cavity in a particular direction be a beam of photons then in a small box inside the cavity we have:

$$I = \frac{h\nu}{\Delta t dA}$$

$$u = \frac{h\nu}{\Delta l dA}d\omega$$

dA is the surface area of the small box with length Δl.
Δt is the average time between photons.
Δl is the average distance between the photons.
I is the intensity of the radiation passing through the box.
u is the energy density in the box with a contribution only from dω.

By integrating over all directions we have:

$$u_\nu(\nu) = I_\nu(\nu)\frac{\Delta t}{\Delta l}\int d\omega = I_\nu(\nu)\frac{4\pi}{c}$$

We then have the radiation intensity out of a pinhole in the cavity given by:

$$I_\nu(\nu) = u_\nu(\nu)\frac{c}{4\pi} = \frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{k_B T}} - 1}$$

The unit is $W/m^2/Hz$.
And as a function of wavelength:

$$I_\lambda(\lambda) = I_\nu(\frac{c}{\lambda})\frac{\frac{c}{\lambda}}{\lambda} = \frac{2hc^2}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1}$$

The unit is $W/m^2/m$.

Energy of the harmonic oscillator

Classical case

The energy of the oscillator is the sum of the kinetic and the potential energy:

$$E(p, x) = \frac{1}{2m}p^2 + \frac{\kappa}{2}x^2$$

From Statistical Mechanics we have:

$$\begin{aligned}
Z &= \int_{-\infty}^{\infty} e^{-\beta\frac{1}{2m}p^2}dp \int_{-\infty}^{\infty} e^{-\beta\frac{\kappa}{2}x^2}dx && (2.2)\\
&= \sqrt{\frac{2\pi m}{\beta}}C1\sqrt{\frac{2\pi}{\kappa\beta}}C1 && (2.3)\\
&= \frac{2\pi}{\beta}\sqrt{\frac{m}{\kappa}}C1C2 && (2.4)
\end{aligned}$$

The average energy is then:

$$E = -\frac{\partial logZ}{\partial\beta} = \frac{1}{\beta} = k_B T$$

Quantum case

From Quantum Mechanics we know the energy of the n'th mode:

$$E_n = n\hbar\omega$$

From Statistical Mechanics we have:

$$Z = \sum e^{-\beta n\hbar\omega} = \sum(e^{-\beta\hbar\omega})^n \approx \frac{1}{1 - e^{-\beta\hbar\omega}}$$

The average energy is then:

$$E = -\frac{1}{Z}\frac{\partial Z}{\partial\beta} = \frac{\hbar\omega e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}} = \frac{h\nu}{e^{\frac{h\nu}{k_b T}} - 1}$$

Chapter 3

The Photoelectric Effect

Albert Einstein read the paper by Max Planck, but was as Max Planck skeptical and didn't believe the theory was a good description of reality. Albert Einstein's attitude changed completely, when he in 1902 received the results of an experiment done by Philipp Lenard (A similar experiment had been done by Heinrich Hertz in 1887). The experiment demonstrated what is now called the photoelectric effect, see Figure 3.1.

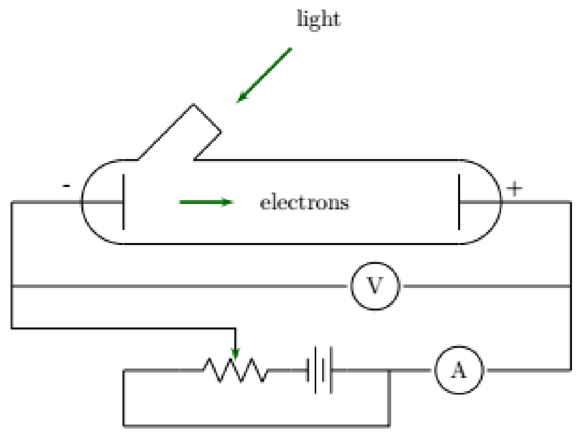

Figure 3.1: Photoelectric Effect setup

In an evacuated glass tube two metal plates are mounted in either end. A voltage is applied between the anode and the cathode. The cathode is exposed to light from a light source through a small window and the current in the electrical circuit is measured.
When the cathode was exposed to light, it was expected that electrons would be removed from the surface and a current could be observed on the ampere meter, if the intensity of the light was sufficiently high..
But a current was only observed if the frequency of the light was higher than a certain threshold frequency, ν_0. Above ν_0 the current was independent of the frequency, but proportional to the intensity of the light. The threshold frequency was seen only to depend on the material used for the cathode, see Figure 3.2.

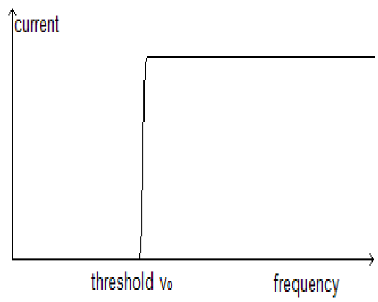

Figure 3.2: Frequency threshold

In 1905 Albert Einstein published a paper, having Max Planck's paper in mind, where he concluded:

The light in this experiment has to be regarded as consisting of particles (photons) having an energy equal to hν, h being the Planck constant and ν is the frequency of the electromagnetic wave.

Physicists were very skeptical about this theory, but when Arthur Compton about 15 years later in 1923 published his paper about what is now known as the Compton effect, it was generally accepted that electromagnetic waves sometimes had to be regarded as particles.

Links:
https://en.wikipedia.org/wiki/Photoelectric_effect
https://byjus.com/jee/photoelectric-effect

Explanation of the findings

This phenomenon can be explained by introducing photons. Specifically, photons with a frequency higher than ν_0 possess an energy exceeding $h\nu_0$. This energy is equal to the energy E_{bindig} required to remove an electron from the surface of the cathode:

$$h\nu_0 = E_{binding} \tag{3.1}$$

If the light intensity is held fixed but the frequency is further increased then the kinetic energy of the electrons is increased:

$$h\nu = E_{binding} + E_{kin} \tag{3.2}$$
$$E_{kin} = h\nu - h\nu_0 \tag{3.3}$$

The current is thus zero for an anode voltage less than $-V_0$ and is increased with the anode voltage. Since the light intensity is held fixed, only a limited number of electrons are removed

from the surface per unit time and the current saturates independently of the frequency. Figure 3.3 shows the current for fixed light intensity for 3 different frequencies above ν_0.
If the intensity of the light is increased, more electrons are removed from the surface and the current increases proportional to the light intensity, see Figure 3.4.

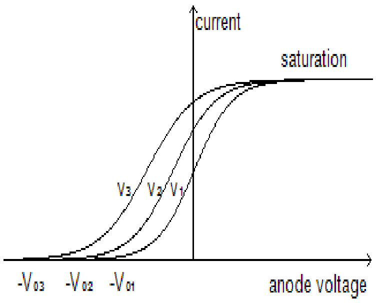

Figure 3.3: Saturation

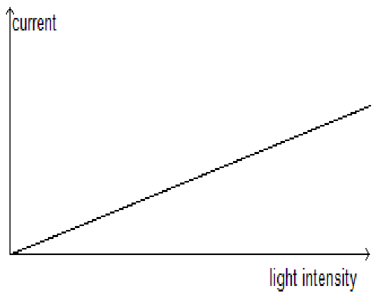

Figure 3.4: Intensity

Chapter 4

The Stern Gerlach Experiment

In 1922 Otto Stern and Walther Gerlach did an experiment with deflection of silver atoms in an inhomogeneous magnetic field.
They were focused on demonstrating space quantization. Space quantization was in those days an idea that stem from the Bohr model of the atom and his quasiplanetary electron orbits and an idea that these orbits may only have discrete orientations in space. Stern's colleagues who had proposed the idea, regarded it as a mathematical construct, more than a physical reality.

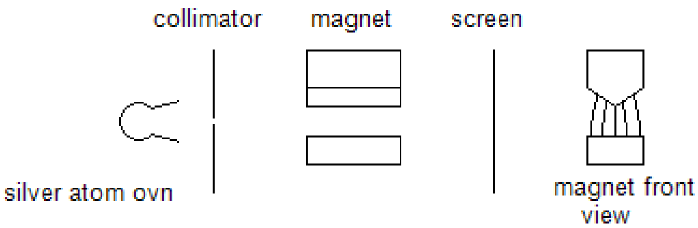

Figure 4.1: Stern-Gerlach experiment

In the experiment sketched in Figure 4.1 silver atoms in the oven to the left are radiated through an inhomogeneous magnetic field, with the field direction being mainly in the z-direction (upwards), but also with a component in the x-direction. If the atoms had a magnetic moment the beam would split according to the direction of the magnetic moment. Stern and Gerlach found that the beam gave rise to two spots on the screen separated in the z-direction. The strength of the magnetic field was 0.1 Tesla and the distance between the two spots was 0.2 mm. Not much, but distinct.
The experiment gave rise to discussions, since the phenomenon was unknown. Some physicist suggested that the result was a consequence of the outermost electron of the silver atom rotated around the nucleus, others that the electron was rotating around itself. Neither explanation was accepted, but finally it was realized that the electron has a property, not seen before. This property is called spin $\frac{1}{2}$. It acts as a sort of intrinsic angular momentum (see the next section). Especially the later work of Paul Dirac on Relativistic Quantum Mechanics(1927), made it possible to theoretically explain the phenomenon.
Links:

https://en.wikipedia.org/wiki/Stern%E2%80%93Gerlach_experiment
https://en.wikipedia.org/wiki/Spin_(physics)
https://www.youtube.com/watch?v=AX9769eQV24

Leonard Susskind and Art Friedman, Quantum Mechanics, The Theoretical Minimum, Basic Book 201.4

More about electron spin

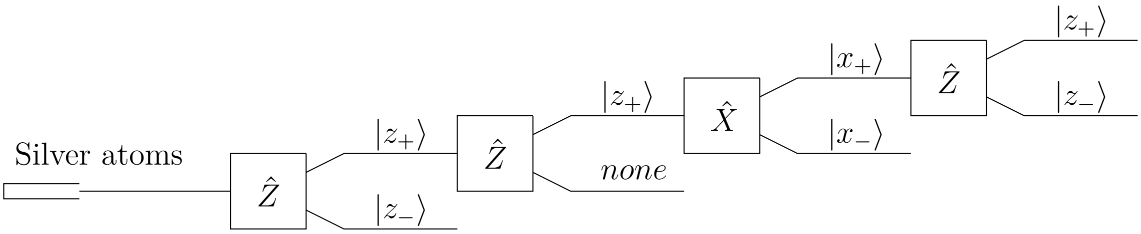

Figure 4.2: Cascading Stern-Gerlach setups

Figure 4.2 shows an experiment (not carried out by Stern and Gerlach) where four setups are cascaded. The boxes labeled $\hat{Z}$ represent a setup with the magnet oriented upwards as above, whereas the box labeled $\hat{X}$ represents a setup with the magnet turned 90 deg. Following the Dirac notation, $|z_+\rangle$ is the state of the atoms deflected up, $|z_-\rangle$ is the state of the atoms deflected down, $|x_+\rangle$ is the state of the atoms deflected to the left and $|x-\rangle$ is the state of the atoms deflected to the right.
The output after the second $\hat{Z}$-box is $|z_+\rangle$ only. This shows that the $|z_+\rangle$ is independent of the $|z_-\rangle$ and that the two states are mutual exclusive and therefore can be regarded as orthogonal. This in turn indicates that the two states should be be described by two orthogonal vectors
After the last $\hat{Z}$-box $|z_-\rangle$ reappear which shows that a $\hat{Z}$-box cannot be viewed as a sorting mechanism of electrons having spin up and electrons having spin down. On the contrary, the electron spin state at the input to the first $\hat{Z}$- box must be viewed as a superposition (linear combination) of the states $|z_+\rangle$ and $|z_-\rangle$.
It is therefore concluded:

Electrons must be viewed as being in a superposition (linear combination) of spin states and there is a probability of 50 % that it is revealed as spin up and 50 % that it is reveiled as spin down. It is not possible to say there is a trajectory for the individual atoms in the Stern-Gerlach experiment.

A no good explantion

As mentioned above the spin has been explained as if the outermost electron of the silver atom rotated around the nucleus or that the electron charge was rotation around the electron itself.

The later does not give much meaning since the electron is regarded as at point particle. How can a point rotate around itself?

Angular momentum of the wavefunction

It was shown by F. J. Belinfante (1939) that the spin can be regarded as an angular momentum generated by a circulating flow of energy in the wave field of the electron. Likewise, the magnetic moment may be regarded as generated by a circulating flow of charge in the wave. The theory gives an explanation similar to the theory of polarized light, which is covered in the next section

Spin in three dimensions and the Hilbert space

Since the two states $|z_+\rangle$ and $|z_-\rangle$, as described above are orthogonal, a two-dimensional vector space with an inner product is required (two vectors are orthogonal when the inner product is zero). Moreover it must be a complex vector space, see below. Such a vector space is called a Hilbert space.
The orthonormal basis is chosen to be:

$$\begin{array}{cc} |z_+\rangle & |z_-\rangle \\ \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{array} \tag{4.1}$$

Due to symmetry $|x_+\rangle$ and $|x_-\rangle$ are also orthogonal and can be represented by two orthogonal vectors.
Looking at the $\hat{Z}$-box to the right, Figure 4.2, it is seen that $|x_+\rangle$ appears as $|z_+\rangle$ and $|z_-\rangle$ (with equal probability).
It is therefore assumed that $|x_+\rangle$ can be written:

$$|x_+\rangle = a\,|z_+\rangle + b\,|z_-\rangle$$

where $|a|^2 = |b|^2 = \frac{1}{2}$, since the probabilities of $|z_+\rangle$ and $|z_-\rangle$ are equal.
Similar for $|x_-\rangle$ we have:

$$|x_-\rangle = c\,|z_+\rangle + d\,|z_-\rangle$$

where $|c|^2 = |d|^2 = \frac{1}{2}$.
The possible values for a, b, c, d are $\pm\frac{1}{\sqrt{2}}$ and $\pm\frac{i}{\sqrt{2}}$. We chose: $a = b = c = \frac{1}{\sqrt{2}}$ and $d = -\frac{1}{\sqrt{2}}$ giving the orthonormal basis:

$$\begin{array}{cc} |x_+\rangle & |x_-\rangle \\ \\ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} & \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{array} \tag{4.2}$$

$|y_+\rangle$ and $|y_-\rangle$ can also be written as linear combination of $|z_+\rangle$ and $|z_-\rangle$ giving the orthonormal basis;

$$\begin{array}{cc} |y_+\rangle & |y_-\rangle \\ \\ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} & \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} \end{array} \tag{4.3}$$

Things we can measure are called measurables or observables and are real quantities. Sometimes we combine two observables in one complex number, but this is only a convinient way of looking at a physical phenomenon. The two examples above, the electron spin and the light polarization, are described using the same mathematics, but the light wave is real and could be described in vector form, but the complex form $e^{ikx-i\omega t}$ is a lot more convenient. The spin state on the other hand is a complex quantity. The spin state itself cannot be measured, only the state along an axis parallel to a magnetic field can be measured. It is therefore necessary to introduce the concept of operators, that can operate on the complex state and at the same time be used to deduce the observables. If the operator is chosen to be hermitian i.e:

$$m_{ij} = m_{ji}^*$$

then the eigenvalues are real an can represent the observables.
If the expected value of a hermitian operator is introduced as:

$$\langle \hat{S} \rangle = \sum \Psi^* \hat{S} \Psi$$

and the eigenvalues are the solutions to:

$$\hat{S}\Psi = a\Psi$$

then Ψ is the eigenstate corresponding to the eigenvalue a and we have the expected value of $\hat{S}$

$$\langle \hat{S} \rangle = \sum \Psi^* \hat{S} \Psi = \sum \Psi^* a \Psi = a \sum \Psi^* \Psi = a$$

The eigenvalues are thus the same as the expected values of the operator.
For the spin a three dimensional operator is introduced:

$$\hat{S} = (\hat{S}_x, \hat{S}_y, \hat{S}_z)$$

For $\hat{S}_z$ the eigenstates are:

$$\begin{array}{cc} |z_+\rangle & |z_-\rangle \\ \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{array} \tag{4.4}$$

and the observables (eigenvalues) are $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$.
The $\hat{S}_z$ operator is then given by:

$$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.5}$$

Correspondingly for $\hat{S}_x$ and $\hat{S}_y$ we have:

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{4.6}$$

$$\hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{4.7}$$

The three matrices:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.8}$$

are called the Pauli matrices and a linear combination together with the unity matrix gives all possible two-dimensional hermite matrices:

$$M = \begin{pmatrix} a+d & b-ic \\ b+ic & a-d \end{pmatrix} + a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + d\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.9}$$

Note that we have:

$$[\hat{S}_z; \hat{S}_x] = i\hbar\hat{S}_y;$$
$$[\hat{S}_x; \hat{S}_y] = i\hbar\hat{S}_z;$$
$$[\hat{S}_y; \hat{S}_z] = i\hbar\hat{S}_x;$$

Polarized light and electron spin

Spin is divided into two types: spin $\pm\frac{1}{2}$ and spin ±1. Photons have spin ±1. The photon spin can be observed as polarization and is mathematically equivalent to the electron spin. Since polarization of light is directly observable and thereby less abstract, an understanding of polarization theory may help understanding the electron spin. However, even both are described using complex math, polarization can be explained without complex math, whereas electron spin cannot.
Light is an electromagnetic wave seen from a classical point of view. It can be polarized in different ways, e.g. horizontally or vertically. A horizontally polarized electromagnetic wave has a horizontal electric field. For vertically polarized light the electric field is vertical. Any kind of polarization can be viewed as a linear combination of horizontal and vertical polarization, assuming the magnitude is 1:

$$\vec{E}_h = e^{ikx-i\omega t}\vec{e}_h$$
$$\vec{e}_h = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$\vec{E}_v = e^{ikx-i\omega t}\vec{e}_v$$
$$\vec{e}_v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$\vec{E} = a\vec{E}_h + b\vec{E}_v = |a|e^{j\phi}\vec{E}_h + |b|e^{j\theta}\vec{E}_v$$
$$\vec{e} = a\vec{e}_h + b\vec{e}_v = |a|e^{j\phi}\vec{e}_h + |b|e^{j\theta}\vec{e}_v$$

ϕ and θ are the phase angles of the electromagnetic wave and $|a|^2 + |b|^2 = 1$.
For polarization turned + or - 45 deg we have:

$$\vec{e}_{+45} = \frac{1}{\sqrt{2}}\vec{e}_h + \frac{1}{\sqrt{2}}\vec{e}_v = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$$

$$\vec{e}_{-45} = \frac{1}{\sqrt{2}}\vec{e}_h - \frac{1}{\sqrt{2}}\vec{e}_v = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$$

For circular polarization (left and right) we have:

$$\vec{e}_{cl} = \frac{1}{\sqrt{2}}\vec{e}_h + e^{\frac{i\pi}{2}}\frac{1}{\sqrt{2}}\vec{e}_v = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$$

$$\vec{e}_{cr} = \frac{1}{\sqrt{2}}\vec{e}_h - e^{\frac{i\pi}{2}}\frac{1}{\sqrt{2}}\vec{e}_v = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$$

Since the $\vec{e}$ vectors are the same as electron spin state vectors we have the following table:

Pol. Type	Spin state	Vector
H	$\lvert z_+\rangle$	$\begin{pmatrix}1\\0\end{pmatrix}$
V	$\lvert z_-\rangle$	$\begin{pmatrix}0\\1\end{pmatrix}$
P_{+45}	$\lvert x_+\rangle$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$
P_{-45}	$\lvert x_-\rangle$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$
P_{cl}	$\lvert y_+\rangle$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$
P_{cr}	$\lvert y_-\rangle$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$

Table 4.1: Cascading Stern-Gerlach setup, state transitions

Figure 4.3 and Figure 4.4 illustrates the orthogonality of vertical and horizontal polarization, In the first all input is seen at the output, in the second there is no output. But by inserting a a polarizer rotated + or - 45 deg it is seen that the vertical polarization can be viewed as a sum of polarizations in these directions. When the third polarizer is turned 90 deg, the light is polarized in the vertical direction. Note that turning the secon polarizer - 45 deg or 135 deg , gives the same result and the directions are equivalent.

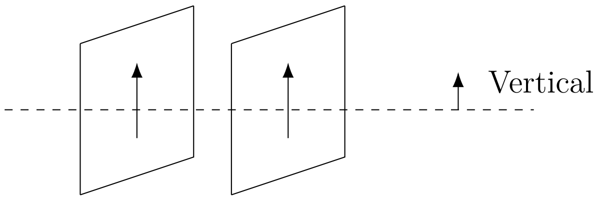

Figure 4.3: Polarization V-V

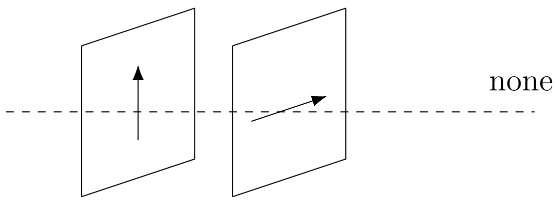

Figure 4.4: Polarization V-H

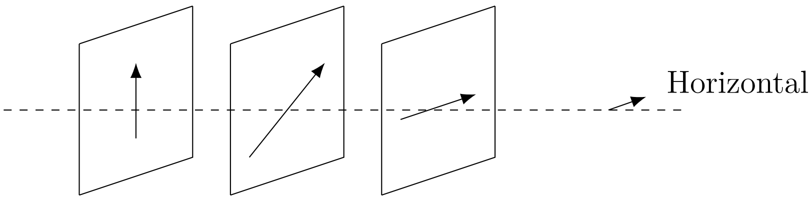

Figure 4.5: Polarization V-VH_45-H

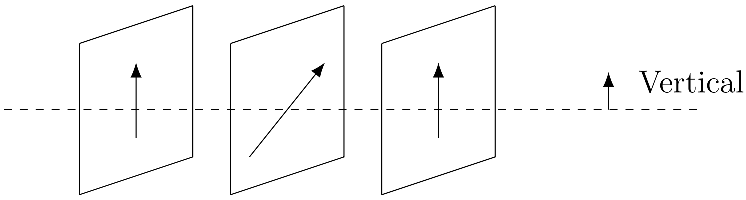

Figure 4.6: Polarization V-VH_45-V

Chapter 5

The Compton Effect

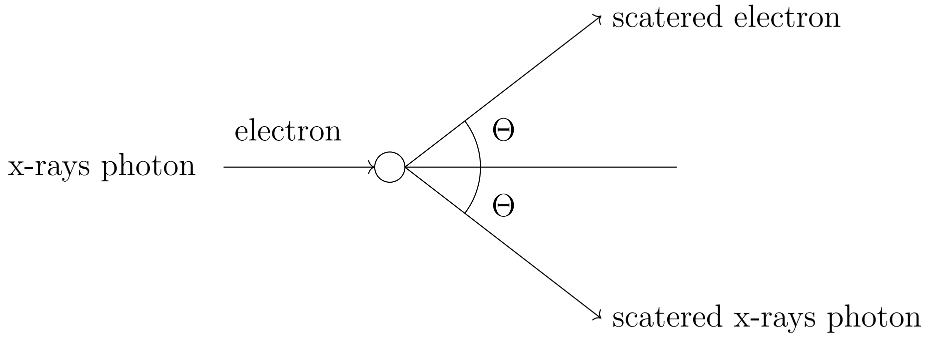

Figure 5.1: Compton setup

I 1923 Arthur Compton carried out an experiment with x-ray scattering, see Figure 5.1. It had been known for some time that the scattered x-rays had a larger wavelength (and thereby lower frequency) than the incoming x-rays. But Arthur Compton was the first to give an explanation for the phenomenon. When x-rays are scattered a recoiling electron can also be observed and by assuming the collision is elastic, it was possible for him to foresee the wavelength of the scattered X-rays. The conclusion was:

X-rays must in this experiment be regarded as particles with energy hν

Just like the finding by Albert Einstein about 15 years earlier. This time physicists were convinced that in some situation electromagnetic waves must be regarded as particles. The experiment led Louis de Broglie in 1924 to suggest the analogy, that particles like protons and electrons also in some situations must be regarded as waves, "Matter waves", with a wavelength of $\lambda = \frac{h}{p}$ called the de Broglie wave length and energy $E = \frac{hc}{\lambda} = h\nu$. This in turn led to the works in 1926 by Werner Heisenberg and Erwin Schrödinger: Quantum mechanics was born.

Derivation of the scattering formula

Looking at Figure 5.1 we have:
The conservation of energy

$$
\begin{aligned}
E_\gamma + E_e &= E_{\gamma'} + E_{e'} \\
E_\gamma &= h\nu \\
E_{\gamma'} &= h\nu' \\
E_e &= m_e c^2 \\
E_{e'}^2 &= (p_{e'}c)^2 + (m_e c^2)^2
\end{aligned}
$$

giving

$$
\begin{aligned}
h\nu + m_e c^2 &= h\nu' + \sqrt{(p_{e'}c)^2 + (m_e c^2)^2} \\
p_{e'}^2 c^2 &= (h\nu - h\nu' + m_e c^2)^2 - m_e^2 c^4
\end{aligned}
$$

Conservation of momentum

$$
\begin{aligned}
\mathbf{p}_\gamma &= \mathbf{p}_{\gamma'} + \mathbf{p}_{e'} \\
\mathbf{p}_{e'} &= \mathbf{p}_\gamma - \mathbf{p}_{\gamma'} \\
p_\gamma &= \frac{h}{\lambda} = \frac{h\nu}{c} \\
p_{\gamma'} &= \frac{h}{\lambda'} = \frac{h\nu'}{c}
\end{aligned}
$$

giving:

$$
\begin{aligned}
p_{e'}^2 &= \mathbf{p}_{e'} \cdot \mathbf{p}_{e'} \\
&= (\mathbf{p}_\gamma - \mathbf{p}_{\gamma'}) \cdot (\mathbf{p}_\gamma - \mathbf{p}_{\gamma'}) \\
&= p_\gamma^2 + p_{\gamma'}^2 - 2p_\gamma p_{\gamma'} cos\theta \\
p_{e'}^2 c^2 &= (h\nu)^2 + (h\nu')^2 - 2h\nu h\nu' \cos\theta
\end{aligned}
$$

We now have:

$$
\begin{aligned}
(h\nu - h\nu' + m_e c^2)^2 - m_e^2 c^4 &= (h\nu)^2 + (h\nu')^2 - 2h\nu h\nu' \cos\theta \\
\nu m_e c^2 - \nu' m_e c^2 &= \nu h\nu'(1 - \cos\theta) \\
\frac{m_e c^2}{\nu'} - \frac{m_e c^2}{\nu} &= h(1 - \cos\theta)
\end{aligned}
$$

And finally we have the observed shift in wavelength:

$$
\lambda' = \lambda + \frac{h}{m_e c}(1 - \cos\theta)
$$

Chapter 6

Electron diffraction

The fact that particles sometimes must be regarded as waves, was first demonstrated by the Thomson-Reid Transmission Diffraction Experiment (1927) and by the Davisson-Germer experiment (1923-1927). As explained above light had been regarded as waves whereas electrons had been regarded as particles and that only light could be diffracted, but the experiments mentioned above showed that electrons indeed can show diffraction patterns. In order to see the diffraction pattern low energy electrons must be used and this requires in turn effective vacuum, since otherwise the electrons are absorbed by the remaining gas.
Effective vacuum had not been possible earlier and that is why the diffraction pattern had not been seen before these experiments. It has later been shown that also protons and neutrons may reveal wave properties.
In case of the Thomson-Reid experiment, the electron beam is directed towards a celluloid film and then detected by a photographic film. The diffraction pattern looks like the picture in Figure 6.1. Given the electrons energy E (approx. 25 keV) the de Broglie wavelength is $\frac{h}{p}$ =7.5pm. Later Thomson alone made similar experiments with aluminum and gold film also showing the diffraction pattern.

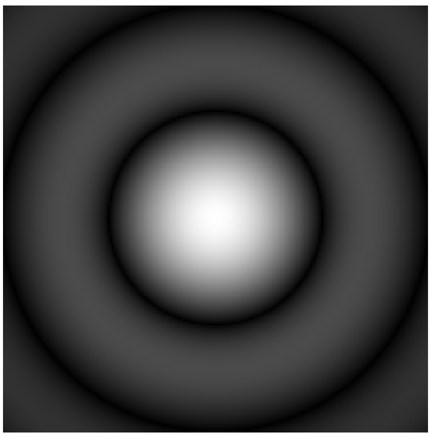

Figure 6.1: Thompson-Reid

In case of the Davisson-Germer experiment, low energy electrons were fired against a nickel crystal and the reflected electron intensity was recorded as a function of angle and it was shown that the

result was similar to X-ray diffraction. After several experiments it was possible to find the de Broglie wavelength.
Links:
https://en.wikipedia.org/wiki/Electron_diffraction
https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1928.0022
https://en.wikipedia.org/wiki/Davisson%E2%80%93Germer_experiment

Chapter 7

The Double-Slit Experiment

The double-slit experiment using light has been studied since the beginning of the 19th century, Thomas Young in 1801. In those days, it was not clear whether light consisted of particles or waves, but during the 19th century physicists became more and more convinced about the wave theory culminating with James Maxwells development of his famous equations. But from the appearance of the Quantum Mechanics, it was clear that electromagnetic waves must be regarded as particles in a number of situations.

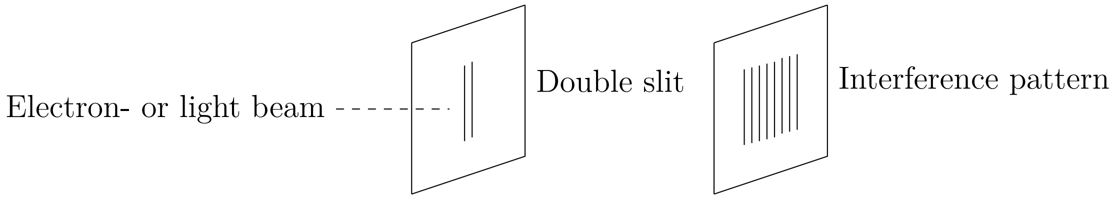

Figure 7.1: Double-slit setup

The double-slit experiment is shown in Figure 7.1. The intension is to demonstrate the wave-particle duality of photons. The setup is as follows: Two closely spaced slits are made in an opaque plate and a light source is placed in front of it. Behind the plate is a screen showing the pattern of what is observed. If one slit is covered the pattern on the screen looks like a small cloud, but if the light can pass through two slits an interference pattern can be seen, see Figure 7.2. If the intensity of the light is lowered an interference is still seen after a while. Even if only one photon is emitted at a time the pattern still shows up, but now after a long time. It is not possible to determine which slit the photon is passing through due to the uncertainty principle derived by Werner Heisenberg.

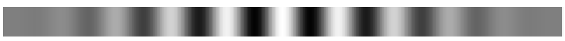

Figure 7.2: Interference pattern

The phenomenon can be explained by stating that the Schrödinger equation has two solutions and a weighted sum is also a solution meaning that the photon enters a superposition which give rise to the photon interferes with itself. See also the Mach-Zehnder interferometer. The

Schrödinger equation and superposition is explained in mere detail in Appendix A. It was not until 1961 the experiment was carried out with electrons by Claus Jönsson. Later others have made similar experiments also demonstrating the wave behaviour of other elementary particles. https://en.wikipedia.org/wiki/Double-slit_experiment

Chapter 8

The Mach-Zehnder Interferometer

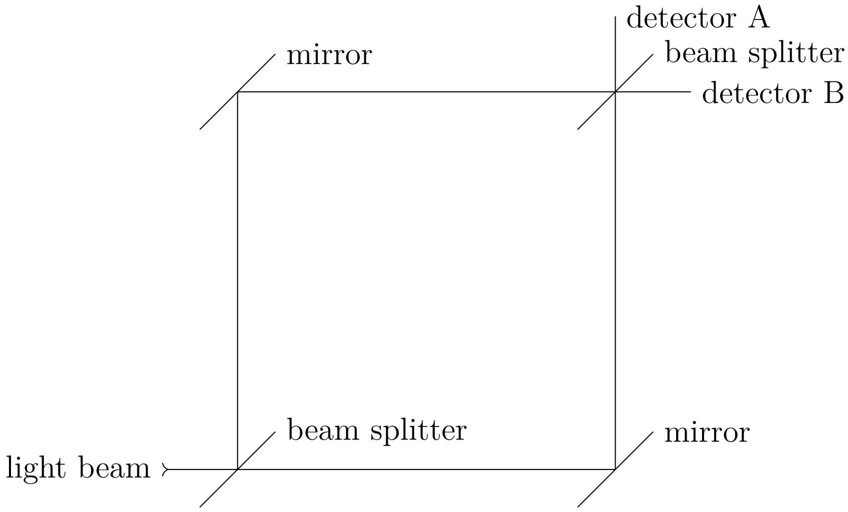

Figure 8.1: Mach-Zehnder setup

The interferometer is shown in Figure 8.1. An incoming light beam to the left is split into two and combined again in the beam splitter to the upper right. Due to the behavior of the beam splitter, it is possible to adjust the mirrors and the two beam splitters so that no light is detected at B and all light is detected at A. The two beams coming from the mirrors are summed and subtracted in the beam splitter to the upper right. If the intensity of the incoming beam is lowered still only detector A receives light.

Even if only one photon is entering the interferometer at a time, only detector A detects the photon. The Schrödinger equation has two solutions for the photon, one going up, the other to the right, at the first beam spliter. The full solution is thus a linear combination of the two solutions. This is called a superposition state. At the beam spliter to the upper right the two solutions are combined and the photon is only detected at detector A.

If the lower vertical path is blocked by a detector C, then this detector detects the photon every second time on the average and every second time on the average the photon is detected by either detector A or B.

The Beam splitter

The beam splitter has been known for more than 200 years. It comes in different designs; the Cube Beam splitter consists of two right angle prisms glued together to form a cube. One of the prisms is coated with a different refractive index. As the name says it can be used to split an incoming light beam into two beams, but can also be used to combine two light beams.

Cassical view

The beam splitter has two inputs and two outputs and the relation between the incoming and outgoing light is:

$$\begin{pmatrix} Y0 \\ Y1 \end{pmatrix} = \begin{pmatrix} h_{00} & h_{01} \\ h_{10} & h_{11} \end{pmatrix} \begin{pmatrix} X0 \\ X1 \end{pmatrix} \tag{8.1}$$

Since the beam splitter is almost loss less, the incoming energy and the outgoing energy must be the same:

$$|h_{00}X_0 + h_{01}X_1|^2 + |h_{10}X_0 + h_{11}X_1|^2 = |X_0|^2 + |X_1|^2$$

giving

$$|h_{00}|^2 + |h_{10}|^2 = 1$$
$$|h_{01}|^2 + |h_{11}|^2 = 1$$
$$(h_{00}h_{01}^* + h_{10}h_{11}^*) = 0$$

Moreover if the beam splitter is balanced, i.e. a single input gives the same intensity for the two outputs we have:

$$|h_{00}X_0|^2 = |h_{10}X_0|^2$$
$$h_{01}X_1|^2 = |h_{11}X1|^2$$

giving

$$|h_{00}|^2 = |h_{01}|^2 = |h_{10}|^2 = |h_{11}|^2 = \frac{1}{2}$$
$$e^{i(\phi_0-\phi_1)} = -e^{i(\phi_2-\phi_3)} = e^{i(\phi_2-\phi_3+\pi)}$$

Examples of ballanced beam splitters are:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & e^{i\frac{\pi}{2}} \\ e^{i\frac{\pi}{2}} & 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & e^{-i\frac{\pi}{2}} \\ e^{i\frac{\pi}{2}} & -1 \end{pmatrix} \tag{8.2}$$

All matrices are unitary:

$$\bar{M}M^* = E$$

Using the beam spliter:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & e^{-i\frac{\pi}{2}} \\ e^{i\frac{\pi}{2}} & -1 \end{pmatrix} \tag{8.3}$$

a corresponding diagram is sown in Figure 8.2.

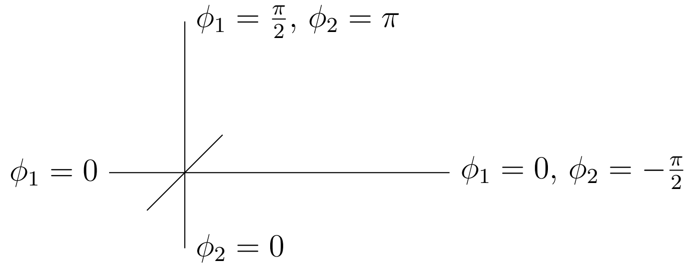

Figure 8.2: Beam spliteter phase

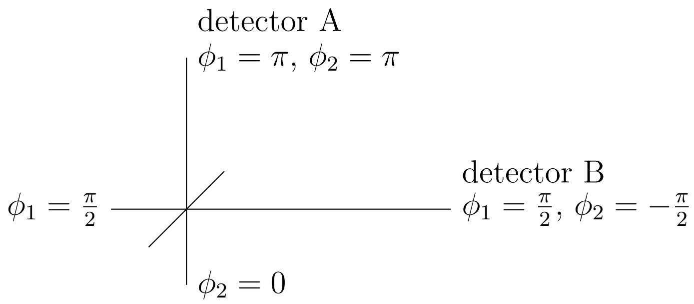

Figure 8.3: Beam splitter II output

From Figure 8.3 it is seen that the outputs at detector B cancels out, whereas the outputs a detector A adds in phase.

Quantum mechanics

If we only look at one photon at a time, a quantum mechanical approach is necessary. There are two solutions to the Schrödinger equation. at the beam splitter. One solution is denoted $|x_u\rangle$ the other $|x_r\rangle$ corresponding to the path up and the path to the right. These two solutons are called states and can also be described by two orthogonal vectors (see Chapter 4):

$$\begin{array}{cc} |x_u\rangle & |x_r\rangle \\ \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{array} \tag{8.4}$$

The superposition state is described by:

$$\alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{8.5}$$

where $|\alpha|^2 + |\beta|^2 = 1$

The sum of the probabities of finding the photon in one of the two paths must be one if the beam splitter is loss less and the two probabilities must be the same if the beam splitter is balanced. The math is thus the same as in the classical case, where energy is replaced by probability.

A single phopton entering the Mach-Zehnder interferometer is therefore always detected at detector A. In case the lower path is blocked there is 50% probability of detecting the photon at detector C , 25% at detector A and 25% at detector B.

Chapter 9

Elitzur-Vaidman bomb

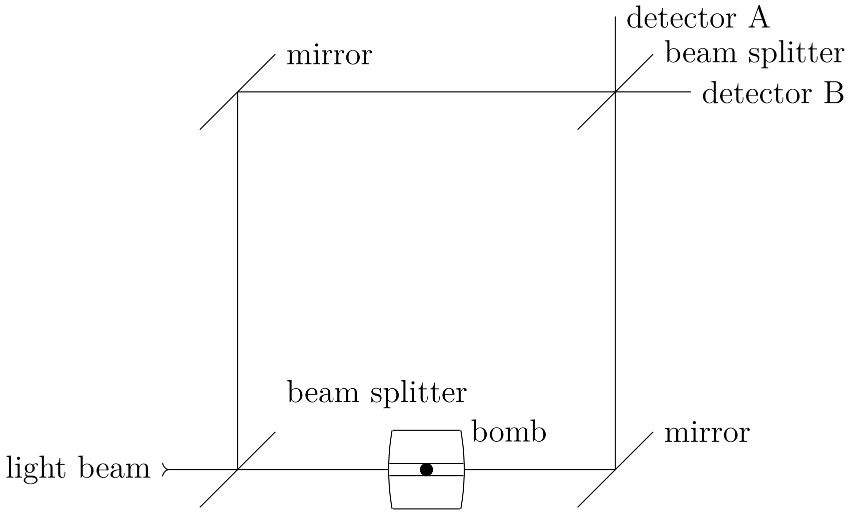

Figure 9.1: Mach-Zehnder setup

The Mach-Zehnder interferometer is shown again in Figure 9.1 now with a bomb inserted in the lower path. The inserted bomb can either be functioning, i.e. if hit by a photon it explodes or the bomb can be defective letting the photon pass unaffected.
As mentioned above if the lower vertical path is unblocked, then a photon is only detected at detector A, whereas if the lower path is blocked by a detector, this detector detects the photon every second time on the average and every second time on the average the photon is detected by either detector A or B.
It it now clear that inserting a functioning bomb corresponds to the situation where the lower path is blocked and the bomb will explode every second time on the average and every second time on the average the bomb does not explode and the photon is detected by either detector A or B.
Inserting a defective bomb corresponds to the situation where the lower path is unblocked and the photon is always detected at detector A. Assuming a collection of bombs where 50% are defective and 50% are functioning, with their states being unknown, the following Table 9.2 shows the possible outcomes on average when a photon is sent to the beam splitter:

Bomb state	Action	Detector A	Detector B
Functioning	Explode	None	None
Functioning	Explode	None	None
Functioning	Survive	None	Detected
Functioning	Survive	Detected	None
Defective	None	Detected	None
Defective	None	Detected	None
Defective	None	Detected	None
Defective	None	Detected	None

Figure 9.2: Elitzur-Vaidman bomb outcome

It is seen that detector B is only detecting the photon if the bomb is functioning an does not explode. It is thus possible to find $\frac{1}{8}$ of the bombs as functioning without destroying them. The cost is that $\frac{1}{4}$ of the bombs are destroyed. $\frac{5}{8}$ are unknown since for these the photon is detected at detector A.

Chapter 10

Entanglement

In 1935 Albert Einstein, Boris Podolsky, and Nathan Rosen published a paper (The so called EPR paper) where they devised a thought experiment showing that according to Quantum Mechanics it should be possible to bring two particles in a combined state with only two outcomes.
If two independent particles can be measured being either in state $|0\rangle$ or $|1\rangle$, then there would be four possible outcomes $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ when measuring the states of both particles. But according to quantum mehanics it is possible to create the two particles in a combined state where only the states $|00\rangle$ and $|11\rangle$ are possible. Or alternatively only the states $|01\rangle$ and $|10\rangle$ are possible. It would then be possible to bring the two particles apart at a large distance and if the state of one of the particles is measured, then the state of the other particle would instantly be known . The phenomenon is now called entanglement.
According to the EPR paper this violates the principles of locality (only local phenomenons can have an effect locally) and that information cannot be transferred faster than the speed of light. Their conclusion was therefore that quantum mechanics was incomplete and that the particles must have properties called ”hidden variables” (some times called supplementary parameters) in order to avoid the ”spooky action at a distance” (information cannot be transferred faster than the speed of light). It is important to recognize that the EPR paper was published without Einsteins approval furthermore Einstein was in no way in opposition to the correctness of quantum mechanics or the reality of entanglement, it was a question of interpretation. Something was missing in order to give a satisfactory explanation of observations. The discussion continued between Albert Einstein and Niels Bohr, who refused to accept that quantum mechanics was incomplete. However, most physicists concentrated more and more on the many questions that could now be answered by quantum mechanics.
But in 1964 John Bell published a paper where he devised a test, Bell's inequalities, that showed that the hidden variable theory was incompatible with quantum mechanics. Clauser and Freedman made experiments in 1972 to test Bell's inequalities and the experiments were not in favor of the hidden variable theory. Later in the early 1980s Alan Aspect finally made several experiments that clearly showed that the hidden variable theory could not be correct since Bell's inequalities were violated and that the experiments were in accordance with quantum mechanics. The consequences were foreseen by Einstein:

Either
We must drop the need of independence of the physical realities in different parts of space.
or
We must accept that measurements of one system can change (instantaneously) the real situation of another system.

So it is a question of whether the two particles should be regarded as one system spread out over space, or two systems instantaneous linked together.

Bell's inequalities

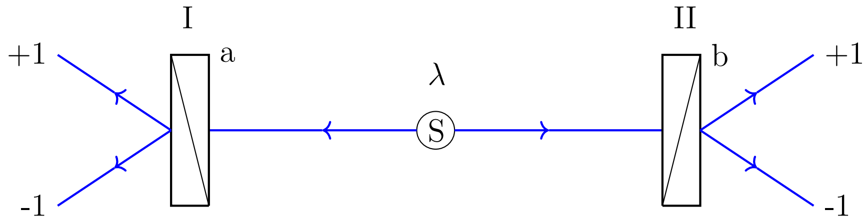

Figure 10.1: Measurement setup

We have an example of two entangled particles when a photon is split into a positron and an electron where the spin of the two particles must be opposite of each other since the total spin must be zero, but the distribution of the spin is random. The probabilities of observing spin up and spin down are each 50%. An other example is two photons being emitted from the same state transition, here the polarization of the two photons are the same, but the distribution of polarization of each photon is random, see Figure 10.1. The result of measuring particle 1 in direction **a** is thus $A(\mathbf{a},\lambda)=\pm1$ an the result of measuring particle 2 in direction **b** is $B(\mathbf{b},\lambda)=\pm1$, where λ represents the initial states of the two particles. If the distribution of λ is $p(\lambda)$ then the expected value of the cross function $A(\mathbf{a},\lambda)B(\mathbf{b},\lambda)$ is :

$$P(\mathbf{a},\mathbf{b})=\int d\lambda p(\lambda)A(\mathbf{a},\lambda)B(\mathbf{b},\lambda)$$

Using a simplified view introduced by Clauser, we consider the quantity:

$$\begin{aligned} s &= A(\mathbf{a},\lambda)B(\mathbf{b},\lambda)-A(\mathbf{a},\lambda)B(\mathbf{b}',\lambda)+A(\mathbf{a}',\lambda)B(\mathbf{b},\lambda)+A(\mathbf{a}',\lambda)B(\mathbf{b}',\lambda) \\ &= A(\mathbf{a},\lambda)[B(\mathbf{b},\lambda)-B(\mathbf{b}',\lambda)]+A(\mathbf{a}',\lambda)[B(\mathbf{b},\lambda)+B(\mathbf{b}',\lambda)] \end{aligned}$$

Since A and B can only take the values ±1 we have:

$$s(\mathbf{a},\mathbf{a}',\mathbf{b},\mathbf{b}',\lambda)=\pm2$$

and the average of s over λ is thus:

$$-2\leq\int d\lambda p(\lambda)s(\mathbf{a},\mathbf{a}',\mathbf{b},\mathbf{b}',\lambda)\leq 2$$

or

$$-2 \leq S(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}') \leq 2$$

where

$$S(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}') = P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{b}') + P(\mathbf{a}', \mathbf{b}) + P(\mathbf{a}', \mathbf{b}')$$

Using Quantum mechanics we have the joint probabilities:

$$p_{++}(\mathbf{a}, \mathbf{b}) = p_{--}(\mathbf{a}, \mathbf{b}) = \frac{1}{2} cos^2(\mathbf{a}, \mathbf{b})$$

$$p_{+-}(\mathbf{a}, \mathbf{b}) = p_{-+}(\mathbf{a}, \mathbf{b}) = \frac{1}{2} sin^2(\mathbf{a}, \mathbf{b})$$

And the cross function is given by:

$$P(\mathbf{a}, \mathbf{b}) = p_{++}(\mathbf{a}, \mathbf{b}) + p_{--}(\mathbf{a}, \mathbf{b}) - p_{+-}(\mathbf{a}, \mathbf{b}) - p_{-+}(\mathbf{a}, \mathbf{b}) = cos(2(\mathbf{a}, \mathbf{b}))$$

and S is given by:

$$S(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}') = cos(\mathbf{a}, \mathbf{b}) - cos(\mathbf{a}, \mathbf{b}') + cos(\mathbf{a}', \mathbf{b}) + cos(\mathbf{a}', \mathbf{b}')$$

Using the vectors in Figure refrefAA2 we have:

$$S_{QM} = S(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}')$$

$$\begin{aligned} S_{QM} &= cos(2\frac{\pi}{8}) - cos(2\frac{3\pi}{8}) + cos(-2\frac{\pi}{8}) + cos(2\frac{\pi}{8}) \\ &= 2\sqrt{2} \end{aligned}$$

It is clear that quantum mechanics violets the Bell inequality.

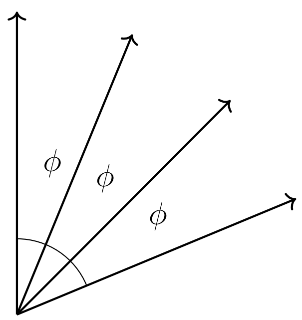

Figure 10.2: Vectors giving max S_{QM}, $\phi = 22.5°$

Experiments by Alain Aspects

Alain Aspect was at an early stage intrigued by the Bell inequality and decided to make an experiment that would show whether the Bell inequalities were fulfilled or not. He devised several experiments, but only the one from 1982 is described here, see Figure Figure 10.3. In the middle

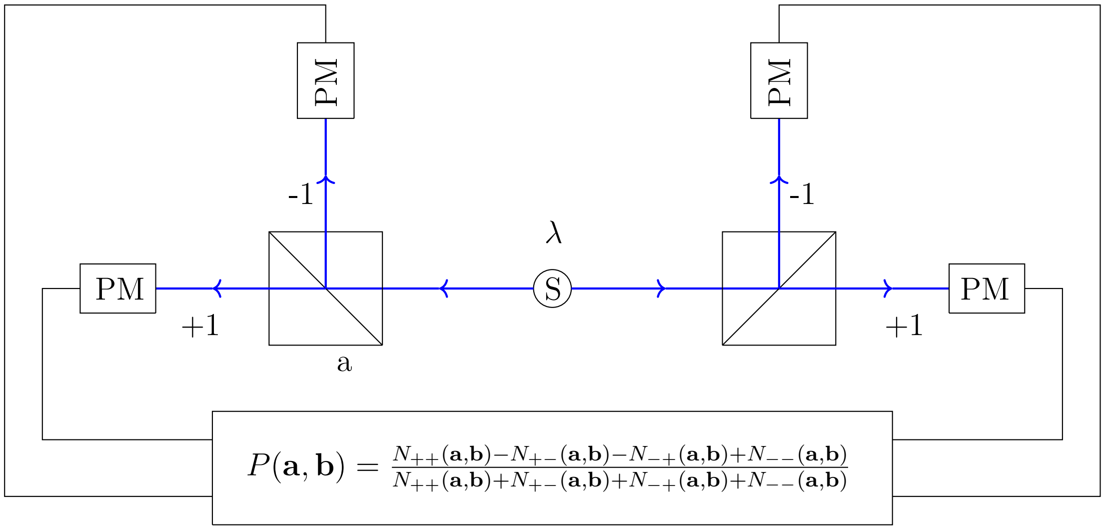

Figure 10.3: 2. Measurement setup

there is a source producing single pairs of entangled photons. A pair is produced every 0.01 sec. The two photons move in opposite directions and hits a polarizer where it is diffracted either up or down corresponding to two orthogonal polarizations. On the figure **a** and **b** are the angles of rotation of the polarizers.
Looking at the photon to the left the output is totally random, the probability of finding the photon in state $+$ is $\frac{1}{2}$ and -1 $\frac{1}{2}$, likewise for the photon to the right. But if $\mathbf{a} = \mathbf{b}$ the joint probability of both photons being instate $+1$ is $\frac{1}{2}$ and -1 is $\frac{1}{2}$, whereas the joint probability of finding one in state $+1$ and the other in state -1 is 0. This means that although the output is random at either end, the two photons are perfectly correlated. The distance between the two polarizers is 6 m.
The number of observations of the two photons are N_{++},N_{+-},N_{-+},N_{--} and the cross function is given by the formula in Figure 10.3. The experimental result of S is:

$$S_{exp} = 2.697 \pm 0.015$$

This shows that the Bell inequality is clearly violated.
Taking the measurement inaccuracies in to account, quantum mechanics gives:

$$S_{QM} = 2.70 \pm 0.05$$

It is therefore concluded that the hiden variable theory is incorrect and that quantum mechanics is in accordance with experiments. We therefore is left with the prediction by Albert Einstein as mentioned in the beginning of this chapter.

Chapter 11

Teleportation and entanglement