

Gerold Gründler

Quantum Phenomena and their interpretation

Astrophysical Institute Neunhof, Nürnberg

www.astrophys-neunhof.de

Version 6.5 , May-13-2026

Like printed media, this electronic book is protected by copyright legislation. You may store the file for your own use. You also may print the book for your own use, but the author does not recommend to do this (see page 5). You may grant other persons access to this book. For this purpose please don't forward the file, but the link to the website mentioned above, so that every reader will always have the current version available. It is forbidden to sell or to lend against payment this book or parts of this book in electronic or in printed form. It is forbidden to integrate this book or parts of this book into other books or other types of publications in electronic or in printed form. It is forbidden to distribute or sell or make available to other persons any modified or shortened versions of this book in electronic or printed form.

Content



How should this e-book be read?	5
1 Classical Physics and Quantum Physics	8
2 Light-Waves	22
2.1 Interference	22
2.2 Polarisation	29
2.3 Thomas Young's Experiment	38
3 Photons	43
3.1 The Photoelectric Effect	43
3.2 The Hypothesis of Light-Quanta	47
3.3 Compton-Scattering	53
3.4 Taylor's Double-Slit Experiment	58
3.5 Single Photons interacting with Beamsplitters	60
3.6 Single Photons in the Interferometer	66
3.7 The Position of a Particle	68
3.8 Relational Properties	71
4 Matter-Waves	75
4.1 de Broglie's Hypothesis	75
4.2 Diffraction of Matterwaves by Crystals	77
4.3 Neutron Diffraction by Single- and Double-Slits	80
4.4 Electron diffraction at single- and double-slits	87
4.5 Diffraction of Atoms and Molecules at the Double-Slit	92

5	Vectors and Projection Amplitudes	102
5.1	Unit Vectors	103
5.2	Two Examples	110
5.3	Measurements in Quantum Theory	118
5.4	How Quantum Theorie was detected	119
6	Entanglement	126
6.1	Spin	127
6.2	Measurement and Reality	132
6.3	EPR = Einstein, Podolski, Rosen	134
6.4	(In-)Divisible Phenomena	147
6.5	Bell's Inequality	150
7	Experiments on Bell's Inequality	158
7.1	Verschränkte Ca-Lumineszenz-Photonen	159
7.2	Verschränkte Be ⁺ Ionen	171
7.3	Verschränkte SPDC-Photonen	183
7.4	Raumartig getrennte Detektoren	190
7.5	Schlussfolgerungen	194
8	Which Way?	201
8.1	Heisenberg's Indeterminacy-Relation	201
8.2	The Path through the Interferometer	206
8.3	A Quantum-Eraser	214
8.4	The interaction-free Light-Barrier	231
9	Interpretations	243
9.1	The Copenhagen Interpretation	243
9.2	John v. Neumann's Collaps	252
9.3	Many Worlds	258
9.4	Decoherence	266
9.5	Many Interpretations	277

10 Quantum Systems of Many Particles	279
10.1 Dinge, Bosonen, Fermionen	279
10.2 Halbleiter-Elektronik	287
10.3 Der Laser	300
References	356

How should this e-book be read?

For several reasons, this e-book should not be printed, but be read from screen¹:

- * At times improved versions are published and posted in the net. A paper print might soon be out-dated.
- * There are countless links in this book. (This is of course exaggerated, they could be counted. . .) It only needs a mouse click to jump to the linked targets, and to jump back subsequently (using the readers “Go to Previous View” button, or the   key-combination) at same speed, instead of cumbersome browsing in a paper print.
- * Sometimes it is useful, to view at the same time several formulas or pictures, which are located on different pages. While the file can be viewed without difficulty in several windows on the screen at the same time, it would be quite expensive to prepare several prints.
- * By saving ink and paper, you can avoid unnecessary pollution of the environment.



The best way to read the book on screen is to use the Two-Page View of the pdf-reader:

- ✓ Make sure that even pages (2,4,6, . . .) are displayed on the left, and odd pages (3,5,7, . . .) on the right. In some pdf-readers you may need to select the option “Show Cover Page”



¹ of course from the screen of a PC or tablet, not from a smartphone display!

- or “Book View” or something similar in the “View / Page Display” settings.
- ✓ Also, make sure that the top and bottom edges of the page are aligned with the top and bottom of the reader’s window, but never in the middle. Otherwise, your eyes will quickly start to water and you’ll lose track of where you are. Therefore, *never* select options like “Enable Scrolling” or “Display Pages Continuously” or anything similar!

Only if your screen is too small to display a full double-page at appropriate view-size (usually 100 % or larger), select single-page view. To do this, select in the “View / Page Display” settings the option

- ✓ “Single Page View” (but not “Single Page Continuous” or “Enable Scrolling” or anything similar, which would be *completely unsuitable!* It is very important that at any time the top or bottom edge of a page is aligned with the top or bottom edge of the window, but not somewhere in the middle. Otherwise, your eyes will quickly start to water and you’ll lose track of where you are.)
- ✓ Set the zoom level to 100 % and resize the reader’s window so that a page is visible in full width and at least 2/3 of its height. Using the  and  keys on the keyboard, you can then scroll with two clicks one page forward or backward. Of each page you will first see the top and then the bottom portion, with about 1/3 page-height overlap. This text view will seem surprisingly familiar to you, because in fact, you unconsciously view the pages of a printed book in much the same way.

You may find the reader’s “Read Mode” convenient, as it provides a larger viewing area by hiding the toolbar. To return to the previous

view in this mode after following a link, press the   keys simultaneously.

1 Classical Physics and Quantum Physics

Since when does thinking exist on Earth? Of course, no one knows for sure, but a rough estimate is possible. Rocks, heads of lettuce, and puddles of water cannot think, because thinking requires a brain. Since when do brains exist on Earth? Let's start with simpler structures: Since when have there been molecules on Earth that can replicate themselves, i. e. pass on genetic information due to replication²? Such molecules, the precursors of the genes in today's living organisms, have likely existed on Earth for about 3.8 billion years.

Indirect evidence of early life dates back to approximately 3.8 to 3.5 billion years ago, such as e. g. certain carbon compounds that are likely synthesized by living organisms, or stromatolites (which are mound-like deposits of sediment particles) that are interpreted as the legacy of microbial populations.

Single-celled organisms with a cell nucleus (known as eukaryotes) exist since about 2 billion years. Multicellular organisms, on the other hand, are around for about 1.2 billion years. Before multicellular organisms appeared, there were certainly no brains on Earth, so no one could think.

It was only about 420 million years ago that plants colonized the continents, closely followed by animals. Until then, all life had existed in water. The gills of our ancestors, who left the water at that time, did not disappear entirely, but found a new purpose as

² replication = re-duplication = doubling without change

hearing organs. A less welcome side effect of this organ recycling is that our ears — due to their close connection to our respiratory organs — are quite susceptible to colds.

Over the long period from 230 million years ago to 65 million years ago, the Earth's landmasses were dominated by dinosaurs. About 65 million years ago, they became largely extinct³ within a short period of time, presumably due to a sudden climate change, perhaps caused by the impact of a large meteorite or by massive volcanic eruptions.

Today's rodents, lagomorphs, shrews, and primates (which include humans) shared common ancestors during the age of the dinosaurs, referred to by paleontologists as euarchontoglires. It was not until the end of the dinosaur age that the evolutionary line of primates diverged from that of the other species. Undoubtedly, the euarchontoglires had brains. Did they think? About what? Probably neither about physics nor about the meaning of life; rather about the next step: Where can I find food? Where can I hide so I won't be eaten by other animals? What could that be moving in the bushes? How far do I have to jump, and where do I have to reach to avoid falling out of the tree?

Between approximately 20 million years ago and 5 million years ago, the human lineage diverged from the lineages of various ape species, most recently from that of chimpanzees. Our ancestors, who lived in Africa about 4 million years ago, are classified by paleontologists as the first humans. About 200 000 years ago, homo sapiens⁴ appeared, the modern (and only surviving to this day) species of humans, to which also we modern humans belong.

The human brain had grown considerably since the days of our ancestors in the dinosaur age. Consequently, people began to

³ among the animals living today, only birds are considered descendants of the dinosaurs

⁴ (Latin) homo sapiens = the reasonable man

think more deeply. Sciences emerged. Science means not only that humans arrive at important insights through observation and careful thoughts, but also that they can share and discuss their insights with others. This requires not only brains but also language — and a fairly sophisticated one at that. Something like the cluck-cluck-cluck of chickens is not enough.

The ability to record and pass on knowledge in writing is certainly very useful, but it is not an indispensable prerequisite for science. Only less than ten thousand years ago humans began to produce written documents. Yet even before that, over many tens of thousands of years, they had been exploring — and explaining to their fellow humans — how to make tools, how to start and extinguish fires, what to sow when and into which soil if one wanted to harvest months later, which herbs to use to treat which illnesses, how to build boats suitable for fishing on sea, how to move heavy loads using levers. That all undoubtedly was science.

A prerequisite for any science is a certain degree of reliability and predictability in processes and events. If I observe today that dry wood is better suited than wet wood for starting a fire, it will not turn out tomorrow that dry wood must first be dipped in water before it can be lit. If a boat is built such that it safely carried the fisherman and his catch back to shore today, then it will do so tomorrow again, and will not suddenly sink due to insufficient load-bearing capacity.

Obviously the events in the world do not occur entirely by chance or in chaotic manner, but rather follow rules known as laws of nature. If one knows these rules, then one can draw from the world's present state conclusions about its past and future states. For example, already in ancient times the laws have been recognized, according to which celestial bodies move at certain speeds along certain paths. A document of ancient astronomy that has survived to this day is the *Almagest* [1], written by Ptolemy

around 150. Based on the current positions of the Sun and the Moon, ancient Greek astronomers were able to calculate by these laws the time of the next lunar eclipse.

In his book “*Philosophiae Naturalis Principia Mathematica*” [2], published in 1687, Isaac Newton (1643–1727) described two laws of nature — the law of motion and the law of gravitation — which, due to their simplicity and universal applicability, by far surpassed all previously discovered laws of nature. With the help of these two laws, it was possible not only to calculate the orbits of celestial bodies but also all mechanical processes on Earth.

Why does Nature follow rules? Couldn't she just — for no reason at all — give the moon a little nudge and cause a lunar eclipse at a time no astronomer has foreseen? Newton was not only a brilliant scientist; he also was a devout Christian. He was convinced that the world was God's creation. Therefore he did not ask whether and why Nature follows rules. She does so, of course, because God not only created the world but also the laws of nature that govern it. But was God himself subject to the laws of nature he had created? Newton's answer was a clear No. In his view, God could not only override the laws of nature, but he had to do so in order to sustain his creation.

From the combination of the law of motion and the law of gravitation, Newton was able to deduce that Sun and Earth move in stable elliptical orbits around their common center of mass. He was only able to perform this calculation, however, under the simplifying assumption that no other celestial bodies disturb this system through gravitational interaction. But in reality, there are numerous disruptive factors in the form of other planets and comets. Therefore, Newton surmised that the solar system was unstable and that God needed constantly to intervene and make adjustments, to keep his creation in balance — in sovereign disregard of the natural laws of motion and gravity discovered by Newton!

Between 1799 and 1823, Pierre Simon Laplace (1749–1827) published the five volumes of his “*Traité de Mécanique Céleste*” (Treatise on Celestial Mechanics). Thanks to the considerable advances in mathematics since Newton’s days, Laplace was able to prove in his *Traité* that Newton’s fears were unfounded, that the disturbance of the planets’ orbital paths by other planets in no way destabilizes the entire system, and that therefore no divine intervention is necessary.⁵ When Napoleon asked where, then, there was still room for God in his system, Laplace is said to have replied⁶ proudly and confidently: “I no longer need that hypothesis.” In 1814, Laplace formulated the conclusion, that inevitably followed from his physical research, as follows:

“An intelligence that, at a given moment, knew all the forces at work in nature as well as the relative positions of the elements that compose them, and that were, moreover, comprehensive enough to subject these given quantities to analysis, would encompass in a single formula the motions of the largest celestial bodies as well as those of the lightest atom; nothing would be uncertain to it, and both the future and the past would lie open before it’s eyes. [...] The regularity that astronomy reveals to us in the motion of

⁵ At least not within the short period of time that was assumed in those years to be the age of the world. The Irish bishop James Ussher (1581–1656) had calculated, based on the data available to him, particularly the Old Testament, that God had created the world in the year 4004 B.C. Newton carefully checked Ussher’s calculation, and corrected the age of the world by 534 years downward. Laplace was not impressed by such calculations. But even the atheistic science of the 19th century, when attempting to estimate the age of the universe, tended to think in terms of a few million years, but not the 13.7 billion years which, according to the current view of cosmologists, have passed since the Big Bang. On a timescale of billions of years, the solar system is *not* stable.

⁶ This quote is not historically verified; *si non è vero, è ben trovato*. (Italian: If it is not true, it is well invented.)

comets is undoubtedly present in all phenomena. The curve described by a simple air- or gas-molecule⁷ is governed just as certainly as the orbits of the planets.” [3]

This was not merely Laplace’s personal opinion. His words accurately and precisely reflect the conclusion reached by physics at that time: what happens today is determined in every detail by causes that lie in the past. Admittedly, no one knows the state of the world at any given moment in all its atomic details, and therefore no one can calculate the state of the world at another moment in all its details. But this is a purely practical problem and does not alter the fact that — if the physics known in Laplace’s time was correct — the actual (though not fully known to any human) state of the world at one point in time completely determines its state at every other point in time.

Consequently, even in the age of the dinosaurs every sentence was already determined, which Laplace in the early nineteenth century would write in his *Traité de Mécanique Céleste*, as was every word of his answer to Napoleon’s question about God’s place in the world. Did Laplace really believe in all seriousness what he wrote there? Most people are convinced that they have free will, that they make countless more or less significant decisions every day, which they could have made differently. Is that merely an illusion? Do we in reality, after laboriously weighing all possibilities, ultimately only decide for those alternatives, which anyway have been predetermined since time immemorial?

Laplace’s “comprehensive intelligence” could impossibly calculate the course of world’s history, if humans could constantly confuse

⁷ One cannot blame Laplace for failing to foresee the astonishing discoveries of twentieth-century quantum physics. But one can certainly expect an intelligent person — especially a scientist — to reflect self-critically on what he knows and what he does not know. If Laplace had done so, he would have wisely remained silent on the subject of “atoms and molecules” instead of putting such outrageous nonsense on paper.

everything due to free-will decisions. If the calculation of the “comprehensive intelligence” gave the result that I will not walk to the bakery today, but will instead, as an exception, ride there by bicycle, then I cannot decide by virtue of free will — after weighing the pros and cons of using the bicycle — to walk there by foot.

Free will and deterministic physics are incompatible with one another.⁸ Either free will is an illusion, or there must be a fundamental flaw in the deterministic physics that Laplace so eloquently described. For most people who haven’t been brainwashed by an education in physics, probably the case is clear: With the determinism, physicists have clearly made a blunder; let them figure out themselves how to get out of that mess.

And I suspect that, in fact, very few physicists at any given time shared Laplace’s belief in a deterministic world. Nevertheless, over a period spanning threehundred years — from the early 17th to the early 20th century — they gathered an impressive body of evidence that all events in this world unfold according to deterministic laws of nature, that what is happening now is unambiguously determined by what happened in the past.

The edifice of deterministic physics — now known as classical physics — collapsed in the early twentieth century. Physics is, after all, an **empirical** science. Every day, experimenters test physical theories down to far-flung details, attempting to uncover any discrepancies between theory and observed reality. And in fact, at the beginning of the twentieth century, a rapidly increasing number of phenomena were observed that clearly went beyond the scope of classical physics. These phenomena were termed

⁸ Of course, this depends on how exactly one defines “free will”. The topic is by no means simple and has occupied numerous philosophers for centuries. Depending on their understanding of “free will”, many of them believe that free will and deterministic physics are, in fact, compatible. For an introduction to the philosophical discussion, the Wikipedia article is a good starting point: https://en.wikipedia.org/wiki/Free_will

quantum phenomena, and the new branch of physics suited to their description was called quantum physics. When one wishes to emphasize the mathematical formalization of physics in particular, one also speaks of classical theory and quantum theory.

In hindsight one can see that the collapse of classical physics began in 1896, when Antoine-Henri Becquerel (1852 — 1908) discovered that uranium salts emit an invisible radiation, which blackens photographic plates even through light-proof paper. The phenomenon was termed radioactivity. Later it was understood that the atomic nuclei of uranium-239 decay into the atomic nuclei of thorium-235 and helium-4. The rays consisting of helium-4 nuclei are called α -rays (pronounced: alpha-rays); they cause the darkening of the photographic plates.

For the following reason, radioactivity could not at all be reconciled with classical physics: On the one hand it is possible to determine the half-life of radioactive material with great precision. For example, if a material has a half-life of ten days, then of one gram of that material after ten days only half a gram will remain, with the rest having decayed in the meantime. If you check again after another ten days, you will find only a quarter of a gram of the material undecayed, and after another ten days, only an eighth of a gram.

That's just like rolling dice. How many times do you have to roll a die to get the result \boxplus ? If the die has six sides, only one of which shows 3, and each of the six sides has an equal probability of landing face-up after a roll, then the probability of rolling \boxplus is $1/6$. The probability that a specific roll will *not* result in \boxplus is $5/6$. The probability that a roll will have any result is one. So the probability of the outcome \boxplus can be written as $\frac{1}{6}$ or as $1 - \frac{5}{6}$. For n tosses, the probability that \boxplus does not appear even once is

$$\underbrace{\frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \dots}_{n \text{ times}} = \left(\frac{5}{6}\right)^n,$$

and the probability of getting at least one \square is

$$1 - \underbrace{\frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \dots}_{n \text{ times}} = 1 - \left(\frac{5}{6}\right)^n .$$

As n gets larger, the probability of rolling at least one \square approaches 1. But no matter how often you roll the die (i. e. no matter how large n is), the probability of rolling at least one \square always remains a tiny bit less than one. Even with an arbitrarily large number of rolls, it is not entirely certain that at least one \square will occur. On the other hand, a \square can also come up with the very first roll; the probability for this is $1/6$. For a very large number of rolls, a precise statement is possible: If you roll the die very often, then in roughly one-sixth of the rolls the result will be \square . But the result of a single roll can impossibly be predicted.

The same is true for radioactive atomic nuclei. For a large number of nuclei, we can determine the half-life with precision. But a single one of these atomic nuclei may decay immediately, even though the half-life is still far from being reached, or it may survive for any length of time. We can only make probabilistic statements about the actual lifetime of a single atomic nucleus, just as with regard to individual outcomes when rolling a die.

The radioactive decay blows deterministic physics knockout. If Laplace's "comprehensive intelligence" cannot calculate when a single radioactive atom will decay, then it can also not calculate when the α -particle emitted during the decay will strike which molecules and ionize or otherwise damage them, nor what further consequences this event will have.

These consequences of Becquerel's discovery were not immediately understood by anybody. But soon similar phenomena came to the attention of physicists — and this time pretty obvious — when experimenters began to systematically investigate the absorption of light by gases. Strong absorption can, for example, mean

that one in a thousand molecules in a gas illuminated by a flash of light absorbs light, while 999 do not. Weak absorption can, for example, mean that one in a million molecules absorbs light, while 999 999 do not. Why do some of the gas molecules absorb light while others do not, even though all molecules are built exactly the same and are exposed to exactly the same flash of light? No reason for this is known, this is simply coincidence.

After a more or less short period of time, the gases re-emit the absorbed light, often with a different color. This is called luminescence. What does “after a more or less short period of time” mean? It means exactly the same as in the radioactive decay of atomic nuclei. When a molecule has absorbed light, it is in an excited state. At some point, the excited molecule decays into a non-excited molecule and the emitted luminescent light. The half-life of the excited state can be, for example, one [nanosecond](#). This means: If this gas is excited by a flash of light, then after one nanosecond half of the excited molecules are still in the excited state, while the other half have already emitted the absorbed energy as luminescence-light. After another nanosecond, $1/4$ of the originally excited molecules are still in the excited state, while $3/4$ of the originally excited molecules have emitted the absorbed energy. After another nanosecond, only $1/8$ of the originally excited molecules still are in the excited state. And so on. It is exactly like radioactive decay: one can specify the average lifetime of the excited state precisely. But how long a single molecule will take to emit the absorbed energy as luminescence-light is a matter of chance; one can only make probabilistic statements about it.

The same is true of chemical reactions. Chemists can precisely determine the average rate at which atoms of types A and B combine to form molecules of type AB under given temperature and pressure conditions. But how many times a specific individual atom of type A must collide with atoms of type B before it finally

combines to form an AB molecule is a matter of chance. In some cases, it can happen in the blink of an eye; in others, it can take a very, very long time. Regarding the speed of a chemical reaction involving individual atoms, one can only make probabilistic statements.

Every time an atom somewhere in the world absorbs or emits energy, every time an atom somewhere in the world reacts chemically with another atom (or fails to do so), Newton's God has the opportunity to intervene in the process and accelerate it, or slow it down, or prevent it entirely, without getting into conflict with any law of nature. Natural science aims to be ideologically neutral, and prefers to speak of chance rather than God. Omnipresent chance puts an end to deterministic physics; it opens up a vast scope that every person can fill at will with ideological and religious content,⁹ and it leaves ample room for free will of humans.

Physics does not prove that humans have free will. It merely came to accept, in the early twentieth century, that free will is not impossible. About the detailed processes by which a human brain weighs the pros and cons of various options of action, then decides for one option, and finally puts that decision into practice, brain research to this day (2017) knows virtually nothing.

But could it not be that actually rules do apply where we believe to see chance at work? Rules that are not yet known by today, but which a more advanced science might uncover in centuries to come? In 1964, it was discovered how one can prove, with full scientific rigor, that such rules do not exist. To prove that such rules "do not exist" means something fundamentally different and far more than merely stating that such rules "are not (yet) known". It means to prove that the world in which we live is in fact *not* deterministically predetermined, that reality is not governed in

⁹ Albert Schweitzer(1875–1965): "Chance is the pseudonym God chooses when he wants to remain incognito."

every detail by the laws of nature, but exhibits a fundamental trait of *irrationality*.

The proof, which is based on a specific type of experiments, has been provided many times since the 1980s. One could be satisfied with the statement: “Physicists have proven that the course of events in this world is not deterministically fixed.” But I think that many people, even if they are not physicists, would still appreciate more detailed informations. It is a central aim of this book to explain that proof to the readers.

The proof is difficult, so reading this book will be a challenge. But the effort is worth it: after all, what is at stake is nothing less than truly understanding a central feature of the world we live in.

While almost everybody welcomed that determinism was abolished, the analysis of quantum phenomena led to a second conclusion which physicists did not like at all: it seemed that many quantum phenomena do not fit into human brains.

This means something other than simply that quantum phenomena clash with traditional ways of thinking (the alleged “common sense”). For thousands of years, there have been countless instances of astonishing scientific discoveries that have forced people to refine or modify their familiar concepts. The ancient Greeks, for example, already knew that the Earth is spherical. Using a trigonometric method, Eratosthenes, director of the University of Alexandria, even measured the size of the globe quite precisely as early as in the third century BCE. This forced people to revise the concepts of “up” and “down”, because our antipodes apparently mean a different (namely exactly the opposite) direction by “down” than we do. In the *Almagest* [1, book 1 chap. 7], Ptolemy mocks those of his contemporaries who doubted that the Earth could rest at the center of the universe, because then, due to its immense weight, it would have to plunge into the depths and pierce through the celestial vault. A more modern example: If someone observes an

event A occurring at one location and another event B occurring at a different location *simultaneously*, then a second person moving in a certain way relative to the first person will observe that A occurs *earlier* than B. And a third person, moving relative to the first in yet another specific way, will observe that A occurs *later* than B. The theory of relativity got its name because Einstein demonstrated that a revision of the concepts “simultaneously”, “earlier”, and “later” was necessary, and that these concepts can only be used without contradiction relative to precisely specified observers.

In contrast, many physicists began to doubt whether the paradoxical quantum phenomena discovered in the twentieth century could be resolved by simply revising some concepts. The Greek word paradox is composed of para = beside, and the word doxa = opinion, view, idea. When a situation appears paradoxical to us, that is usually because our information about the situation is incomplete or incorrect, or because we are viewing the situation from an unsuitable conceptual framework. In any case, such paradoxa can be resolved by obtaining the missing information or by adopting a more appropriate perspective.

But could it not be that we are, in principle, unable to resolve the paradoxa we encounter in many quantum phenomena because doing so would require us to change not only the content of our thinking but also our very way of thinking? The way we think is determined by the structure of our brains, and this structure is in turn the result of hundreds of millions of years of our evolution, which (for precisely this reason) was mentioned earlier. To put it in the language of computer technology: We can change the software running in our brains, but we cannot change the hardware. The hardware — that is, the structure of our brains — has evolved in such a way that our ancestors, in every stage of evolution, were able to find and exploit ecological niches in which they could survive

and develop. The ability to understand quantum phenomena would have been a completely superfluous luxury throughout all those millions of years.

Therefore it is not surprising that evolution made no effort whatsoever to equip the physicists, who began researching quantum phenomena in the twentieth century, with a brain hardware suitable for this endeavor. In contrast to the merely apparent paradoxes, which can be resolved with help of adapted concepts and perspectives, a paradox that humans cannot resolve in principle due to their inherited brain structure could be described as an *objective paradox*. If one regards the paradoxa of quantum phenomena as objective paradoxa, then one will not fight against them like Don Quixote against windmill blades, but will instead try to come to terms with the inevitable paradoxa in the most intelligent and bearable way possible. Or is it “just” a software-problem after all? Have we simply not tried hard enough yet? Can the paradoxa of quantum phenomena still be resolved if we only dedicate ourselves to that task with sufficient intelligence and perseverance? Physicists do not agree on this point, and the author of this book is not sure either.

One of the most astonishing paradoxa, which we will encounter repeatedly in this book, is the one-particle-interference. To understand what this is all about, we first need to know what interference is. I will explain that in the following chapter.

2 Light-Waves

2.1 Interference

Figure 2.1 shows on the left a glass prism, as is commonly used in optical devices. The top sketch shows the prism viewed vertically from above, while the bottom sketch shows it viewed from an angle above. The red light beam enters perpendicularly through one side of the prism, is **reflected** at the back, and exits perpendicularly through the other side.

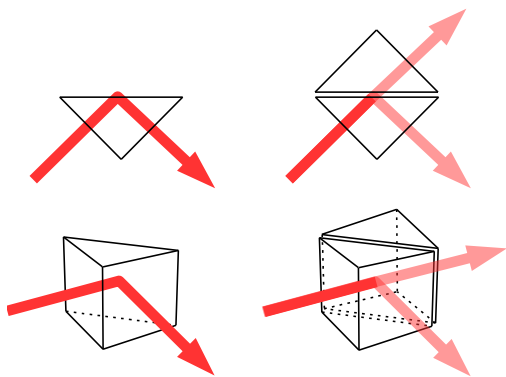


Fig. 2.1: A prism (left) and a beam splitter (right)

When a light-beam, coming from outside, strikes the entrance face of a glass prism, a few percent of the light are reflected. This is known as “external reflection”. Due to appropriate coatings, external reflection can be reduced to nearly zero.

The reflection at the back of the prism is total: 100% of the

light is reflected. This reflection is called “internal reflection”, because the light strikes the reflective surface from inside the glass. The name is somewhat misleading, however, because the reflection does not take place inside the glass; rather, the light travels out of the glass by about one [micrometer](#) before returning into the glass. This becomes visible if two prisms are placed back-to-back. Depending on how flat the back surfaces are ground, a greater or lesser portion of the light no longer returns to the first prism, but enters the second prism and continues to travel in a straight line. This is called “frustrated total internal reflection”.¹⁰ It is used to manufacture beam splitters.

In the right sketch in fig. 2.1, a beam splitter is shown. The top sketch shows the beam splitter viewed vertically from above, while the bottom sketch shows it viewed from an angle above. Beam splitters are usually designed so that 50% of the incident light is [reflected](#) and 50% is transmitted. To achieve this, the two prisms are glued together with a suitable resin so that the distance between their back surfaces is approximately $1\ \mu\text{m}$. In the sketch, the distance is greatly exaggerated. A $1\ \mu\text{m}$ -wide gap can be perceived with the naked eye only as a crack in the glass cube, and even with a good optical microscope it is hardly discernible. With an electron microscope, however, it is clearly visible.

Using prisms and beam splitters, the interferometer shown in figure 2.2 on the following page can be constructed. An interferometer is a measuring instrument used to observe interference. We will explain in a moment what interference is.

The red light beam, coming from the left, strikes the beam splitter BS1, where it is split into the partial beams A and B. Each of the two partial beams is totally reflected by 5 prisms and then strikes the beam splitter BS2. If the length of the paths of both

¹⁰ Physicists can find a detailed description of evanescent electromagnetic fields and frustrated total internal reflection in [4].

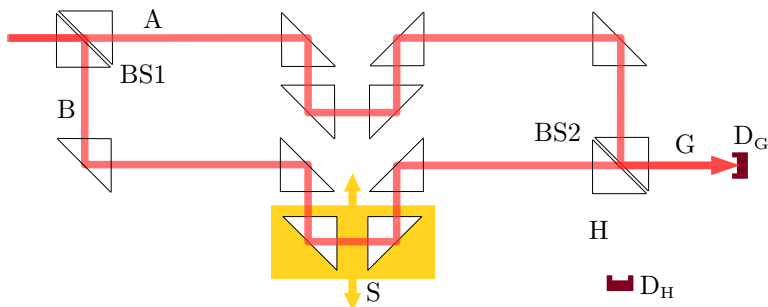
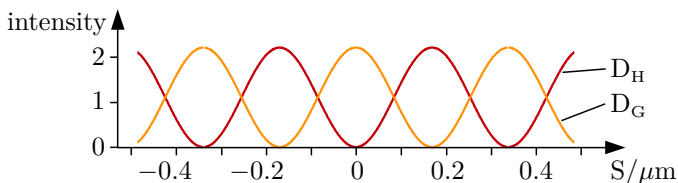


Fig. 2.2: An interferometer

partial beams is exactly the same, then — as drawn in the sketch — all of the light exits at the output G of this beam splitter and reaches the detector D_G . Why this is the case will be explained shortly.

Two of the prisms are mounted on a yellow-outlined slide S . By moving the slide S , the path of light B can be shortened or lengthened. Figure 2.3 shows how the light intensity detected by the detectors D_G and D_H changes when the slide is moved.

As the light path B changes, an increasingly larger portion of the light intensity is detected by the detector D_H , and a correspondingly smaller portion by the detector D_G . If the slide is moved by plus or minus $0.17\ \mu\text{m}$ — i. e. if the light path B is changed by plus or minus $0.34\ \mu\text{m}$ — then all the light reaches the detector D_H ,

Fig. 2.3: Light intensity at D_G and D_H

and no light at all reaches the detector D_G . If the slide is moved even further, more light reaches detector D_G again, and less light reaches detector D_H . If the light path B is changed by plus or minus $2 \cdot 0.34 \mu\text{m} = 0.68 \mu\text{m}$, then the initial state is reached again: All light reaches detector D_G , and no light reaches detector D_H .

How can that be explained? At the beam splitters BS1 and BS2, half of the incident light is transmitted, and half is **reflected**. In total, therefore, $1/4$ of the light travels from path A to path G, $1/4$ of the light travels from path B to path G, $1/4$ of the light travels from path A to path H, and $1/4$ of the light travels from path B to path H. Apparently, light plus light does not always result in more light, but sometimes in less light or even darkness. Physicists have come up with only one explanation for this: light must be a wave. (Further investigation revealed that light is an electromagnetic wave; we will not discuss here the specific basis for that conclusion.)

Why the wave nature of light has the effect that light plus light can sometimes result in darkness, can be read from figure 2.4 on the next page. In each of the three diagrams, the electric field strengths E of wave_A (which previously traveled along path A) and wave_B (which previously traveled along path B) are shown at a specific point in time along path G or H.¹¹ Also plotted is the wave resulting from the superposition of wave_A and wave_B.

The wavelength is defined as the distance between two points of the same “phase”, for example, the distance between two maxima or between two minima of a wave. In 2.4(a) the wavelength is denoted by the letter λ . Because they originate from the splitting of the same wave at beam splitter BS1, wave_A and wave_B (and also their sum) have the same wavelength. Wave_A and wave_B

¹¹ More precisely: The electric field strength divided by Volt is shown. The letter V stands for volts. Electric field strengths are measured in Volt, just like the voltage in the power grid. The field strength of wave_A and wave_B thus oscillates between approximately $+2V$ and $-2V$.

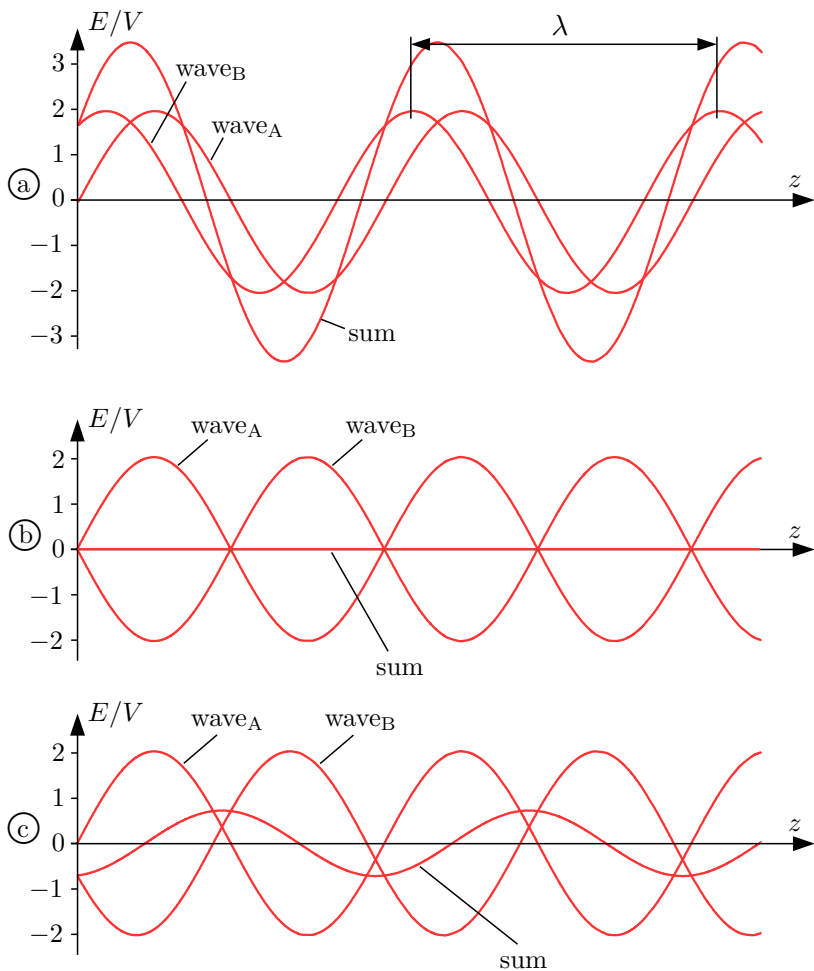


Fig. 2.4: Interference of two waves

also have the same **amplitude**, namely about 2 volts, because the beam splitters transmit and reflect half of the light in each case. But when they are superimposed again on paths G and H, they generally no longer have the same phase, i. e. the maxima and minima of wave_A and wave_B are shifted relative to each other.

The phase shift can have two different causes: Either wave_A and wave_B have traveled different distances because the slide was moved, and/or wave_A and wave_B have been reflected a different number of times.

If the crests and troughs of the two partial waves are located almost at the same points, they add together to form a wave with a large amplitude, as shown in 2.4(a). This is referred to as “constructive interference”. If, on the other hand, the phase difference is approximately half a wavelength, as shown in 2.4(c), then wave_A and wave_B interfere destructively, i. e. the amplitude of their sum is very small. In the extreme case of a phase shift of exactly half a wavelength, wave_A and wave_B completely cancel each other out; their sum results in darkness, see 2.4(b).

The five total reflections each at the back faces of the prisms produce the same phase shift for both partial beams; thus, they do not result in an overall phase difference between the partial beams. At each beam splitter, however, there is an additional phase shift of one-quarter of a wavelength between the transmitted and reflected waves. The reason for this phase shift is irrelevant to our considerations.¹² Light that enters path G via path A was transmitted at the first beam splitter and reflected at the second beam splitter. Light that enters path G via path B was reflected at the first beam splitter and transmitted at the second beam splitter. Overall, therefore, there is no phase shift in path G due to the beam splitters. Thus, the partial waves in path G interfere constructively when the lengths of paths A and B are exactly same.

¹² Physicists will find an interesting study in [5].

The situation is different in path H: Light that enters path H via path A has been transmitted by both beam splitters. Light that enters path H via path B has been reflected by both beam splitters. Overall, there is therefore a phase shift of half a wavelength between the partial beams in path H due to the two beam splitters, and consequently destructive interference. This is why all the light reaches detector D_G , and detector D_H remains dark, if the lengths of paths A and B are exactly identical.

The device shown in figure 2.2 on page 24 is called interferometer, because it can be used to measure the interference of light waves. From diagram 2.3, we can see that the slide must be moved by $0.34 \mu\text{m}$ to get from one position of maximum constructive interference to the next position of maximum constructive interference at the same detector. In doing so, the path length B changes by $0.68 \mu\text{m}$. Therefore, the red light used here has a wavelength of $\lambda = 0.68 \mu\text{m}$.

Humans can perceive electromagnetic radiation with wavelengths between approximately $\lambda = 0.38 \mu\text{m}$ and $\lambda = 0.78 \mu\text{m}$ as visible light. But interference experiments can also be conducted (using differently designed apparatus) with invisible electromagnetic radiation in a wavelength range from approximately 10^{-10}m to about 10 km. Table 2.1 contains a list of electromagnetic radiation with different wavelengths.

The frequency of a wave is the number of wave crests (or troughs) that pass a given point in space per unit of time. Frequency is usually denoted by the letter ν . Frequency is related to wavelength as follows:

$$\nu = \frac{c}{\lambda} = \frac{\text{velocity of the wave}}{\text{wavelength}} \quad (2.1)$$

$c \approx 3 \cdot 10^8$ meter/second is the velocity of light in vacuum or in air. The frequencies of electromagnetic waves are as well displayed in table 2.1. The unit $\text{Hz} = \text{s}^{-1} = (\text{per second})$ is spoken “Hertz”.

	wavelength in air resp. vacuum	frequency
gamma-rays	$< 10^{-12}\text{m}$	$> 3 \cdot 10^{20}\text{Hz}$
X-rays	$10^{-12}\text{m} \dots 10^{-8}\text{m}$	$3 \cdot 10^{20}\text{Hz} \dots 3 \cdot 10^{16}\text{Hz}$
ultraviolet	$10^{-8}\text{m} \dots 0.38 \mu\text{m}$	$3 \cdot 10^{16}\text{Hz} \dots 7.9 \cdot 10^{14}\text{Hz}$
violet	$0.38 \mu\text{m} \dots 0.42 \mu\text{m}$	$7.9 \cdot 10^{14}\text{Hz} \dots 7.1 \cdot 10^{14}\text{Hz}$
blue	$0.42 \mu\text{m} \dots 0.49 \mu\text{m}$	$7.1 \cdot 10^{14}\text{Hz} \dots 6.1 \cdot 10^{14}\text{Hz}$
green	$0.49 \mu\text{m} \dots 0.57 \mu\text{m}$	$6.1 \cdot 10^{14}\text{Hz} \dots 5.3 \cdot 10^{14}\text{Hz}$
yellow	$0.57 \mu\text{m} \dots 0.59 \mu\text{m}$	$5.3 \cdot 10^{14}\text{Hz} \dots 5.1 \cdot 10^{14}\text{Hz}$
orange	$0.59 \mu\text{m} \dots 0.65 \mu\text{m}$	$5.1 \cdot 10^{14}\text{Hz} \dots 4.6 \cdot 10^{14}\text{Hz}$
red	$0.65 \mu\text{m} \dots 0.75 \mu\text{m}$	$4.6 \cdot 10^{14}\text{Hz} \dots 4 \cdot 10^{14}\text{Hz}$
infrared	$0.75 \mu\text{m} \dots 1 \text{mm}$	$4 \cdot 10^{14}\text{Hz} \dots 3 \cdot 10^{11}\text{Hz}$
micro-waves	$1 \text{mm} \dots 1 \text{m}$	$3 \cdot 10^{11}\text{Hz} \dots 3 \cdot 10^8\text{Hz}$
radio-waves	$1 \text{m} \dots 10 \text{km}$	$3 \cdot 10^8\text{Hz} \dots 30 \text{kHz}$
low-frequency	$>10 \text{km}$	$< 30 \text{kHz}$

Tab. 2.1 : Electromagnetic Waves

2.2 Polarisation

We need one more piece of information about light waves: Are light waves longitudinal or transverse waves? What this means can be seen in figure 2.5 on the following page.

In the sketches 2.5(a) and 2.5(b) a rope is displayed whose left end is moved up and down in x -direction so rapidly that a wave propagates along the rope in z -direction. And in sketch 2.5(c) the yellow disc at the left end of a tube is moved back and forth in z -direction so quickly that a sound wave propagates through the tube in z -direction.

In all three sketches, the waves propagate in z -direction. The parts of the rope move up and down in x -direction — i. e. transversally to the direction of wave propagation. Such waves are called *transverse* waves. The air molecules of the sound wave move back and forth in the z -direction — i. e. along the direction of wave

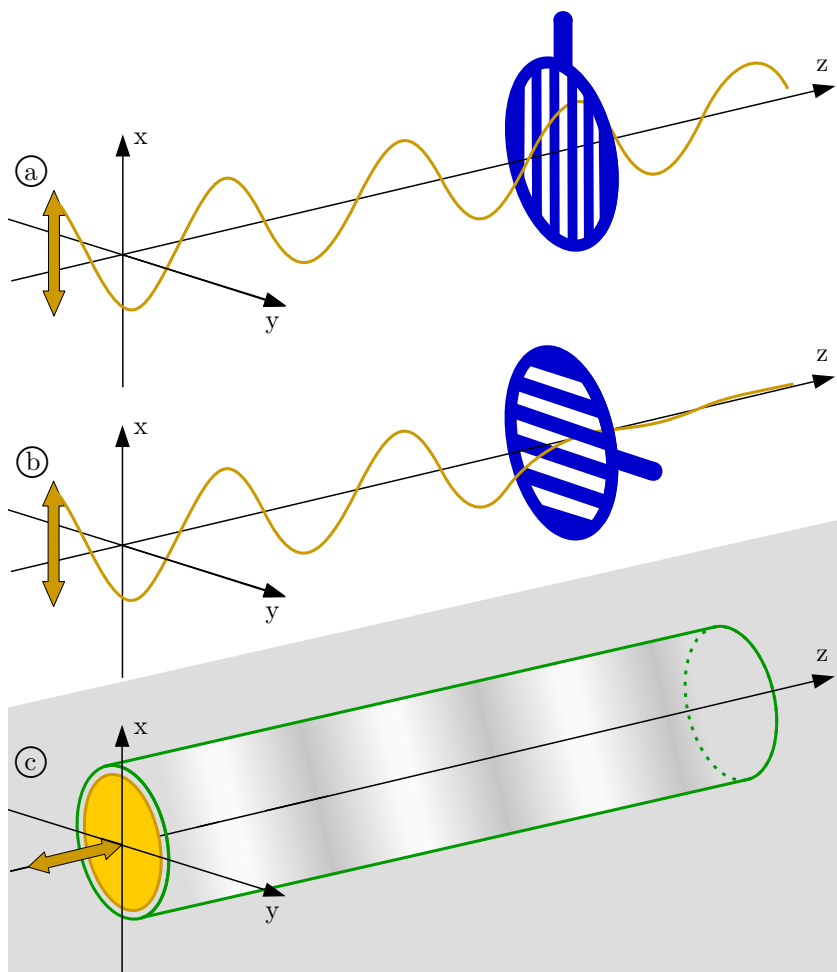


Fig. 2.5: Transverse and longitudinal waves

propagation. Such waves are called *longitudinal* waves.

A characteristic feature of transverse waves is that they can be attenuated to varying degrees due to different rotations of polarization filters. A polarization filter for the rope wave is simply a grid, as shown in blue in fig.2.5(a) and 2.5(b). The parts of the rope oscillate up and down in x -direction. We say that the rope wave is *polarized* in x -direction. In sketch 2.5(a), the grid of the polarization filter is aligned in the x -direction, so that the wave can pass through unimpeded. If the polarization filter is rotated by 90° , the transverse wave is almost completely suppressed, see 2.5(b). In contrast, there exists no filter for the longitudinal wave 2.5(c) whose rotation would have any effect on the degree of attenuation.

Light waves (resp. electromagnetic waves of any wavelength) are transverse waves, because their intensity can be attenuated to varying degrees by rotations of polarizing filters.¹³ In figure 2.6, the red light beam from a lamp is directed through two or three polarization filters onto a white screen.

In all six diagrams, the first filter near the lamp is oriented vertically, so it transmits vertically polarized light. If, as in sketch 2.6(a), the second filter is also oriented vertically, then the maximum amount of light is transmitted to the screen. If, on the other hand, the second filter is oriented horizontally, as in sketch 2.6(c), then no light reaches the screen. The sketches 2.6(a) and 2.6(c) apparently correspond to the sketches 2.5(a) and 2.5(b).

In sketch 2.6(b), the second filter is rotated by 45° relative to the first filter. In this case, half as much light reaches the screen as in the case 2.6(a). The simple mechanical model of a rope and a grid in fig.2.5 does not seem to fit to this result very well, and

¹³ Many commercially available sunglasses are polarizing filters. If you have a pair of such sunglasses at hand, you should definitely try out the amazing effect shown in figure 2.6(f)! A lamp is not necessary; it is sufficient to look at a bright surface through three sunglasses which are rotated relative to one another.

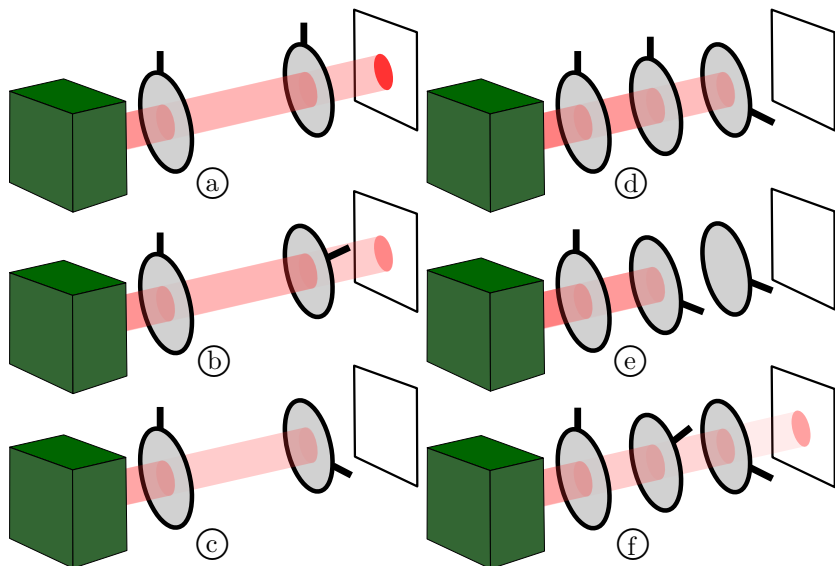


Fig. 2.6: Polarizing filters

even less for the observation depicted in sketch 2.6 (f).

What happens in case of 2.6 (f)? From 2.6 (c), 2.6 (f) differs only in that an additional filter set to 45° has been inserted between the two outer filters set to 90° and 0° . The filters work by absorbing a portion of the incident light, i. e. converting it into heat. Under no circumstances can a filter generate additional light. And yet, surprisingly, the additional filter in 2.6 (f) causes more light to reach the screen than in 2.6 (c). How is that possible?

This is because the analogy with the rope wave in figures 2.5 (a) and 2.5 (b) does not really hold true. Unlike a rope wave, a light wave with arbitrary polarization can always be viewed as the sum of two partial waves, which are polarized perpendicular to each other, as shown in figure 2.7. In 2.7 (a) one looks in the negative y -direction at an electromagnetic wave that propagates in z -direction

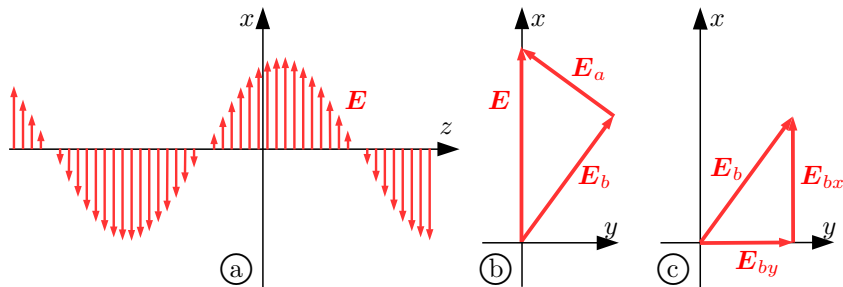


Fig. 2.7: $\mathbf{E} = \mathbf{E}_a + \mathbf{E}_b$, $\mathbf{E}_b = \mathbf{E}_{by} + \mathbf{E}_{bx}$

and is polarized in x -direction. The red arrows symbolize the electric field intensity \mathbf{E} at a specific time at various points in space. In 2.7(b) and 2.7(c) one looks in z -direction onto the xy -plane.

Fig. 2.7(b) illustrates one of the infinitely many ways to represent the field strength \mathbf{E} as the sum of two mutually perpendicular field strengths \mathbf{E}_a and \mathbf{E}_b .

$$\mathbf{E} = \mathbf{E}_a + \mathbf{E}_b ,$$

because one travels from the tail of the arrow representing the electric field \mathbf{E} to its tip, regardless of whether one moves along the arrow \mathbf{E} or sequentially along the arrows \mathbf{E}_b and \mathbf{E}_a .

When the electromagnetic wave \mathbf{E} , which is polarized in x -direction, strikes a polarization filter whose transmission axis is aligned along \mathbf{E}_b , then \mathbf{E}_a is largely absorbed by the filter (i. e. converted into heat), while \mathbf{E}_b is significantly less attenuated. And when the wave \mathbf{E}_b subsequently encounters another polarization filter whose transmission direction is aligned along the y -direction, then \mathbf{E}_{bx} is largely absorbed, while \mathbf{E}_{by} is significantly less attenuated.

While polarizing filters of that type, which is used in sunglasses, are readily available and relatively inexpensive, it is a disadvantage

in our investigations that they absorb a significant portion of the light regardless of the polarization direction, with the absorption merely being particularly strong for light polarized in one direction, but much weaker for light polarized perpendicular to that direction. There exist much better polarizing filters that operate with virtually no loss. They are made from suitably cut and oriented **anisotropic** crystals, e. g. calcite (CaCO_3). In anisotropic crystals, light propagates at different speeds in different directions, and this speed also depends on the polarization of the light.

Using anisotropic crystals, polarizing beam splitters can be fabricated that transmit more than 99 % of light with a specific polarization direction, that **reflect** more than 99 % of light with the perpendicular polarization direction, and that absorb far less than 1 % of the incident light. For simplicity, we will assume in the following that we are working with ideal polarizing beam splitters that transmit 100 % of light with one polarization direction, reflect 100 % of light with the perpendicular polarization direction, and absorb no light at all.

Figure 2.8 on the next page shows the decomposition of light with polarization \mathbf{E} into the mutually perpendicular components \mathbf{E}_y and \mathbf{E}_x . The light enters parallel to the z -axis. The bases of the three beam splitters are oriented parallel to the y - z -plane. The first beam splitter reflects the portion of the light that is polarized in x -direction and transmits the portion that is polarized in y -direction. The other two beam splitters confirm that the light behind the first beam splitter is indeed 100 % polarized in x -direction resp. 100 % in y -direction.

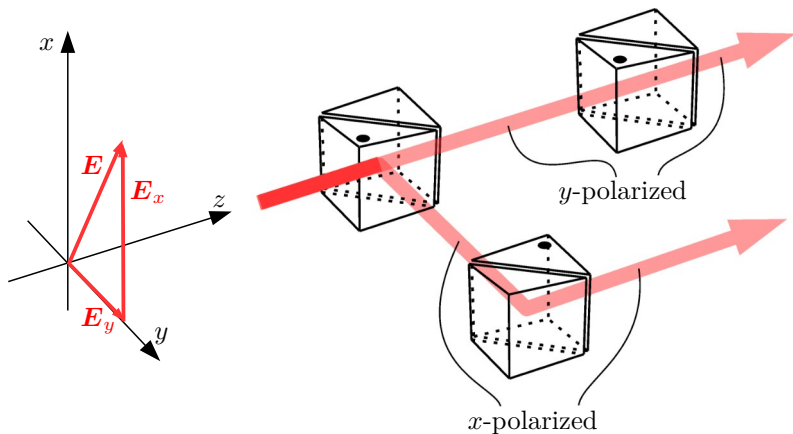


Fig. 2.8: Polarizing beam-splitters

Thus polarizing beam splitters work quite differently from the simple glass beam splitters used in the interferometer fig. 2.2 on page 24. Simple beam splitters transmit and reflect 50% of the incident light, regardless of the light's polarization. In contrast, the polarizing beam splitters shown in fig. 2.8 transmit all light which is polarized in y -direction, and reflect all light which is polarized in x -direction. To prevent that polarizing beam splitters and the so similar looking simple, non-polarizing beam splitters are confused in the laboratory, the manufacturers mark the polarizing beam splitters with a thick black dot on the top.

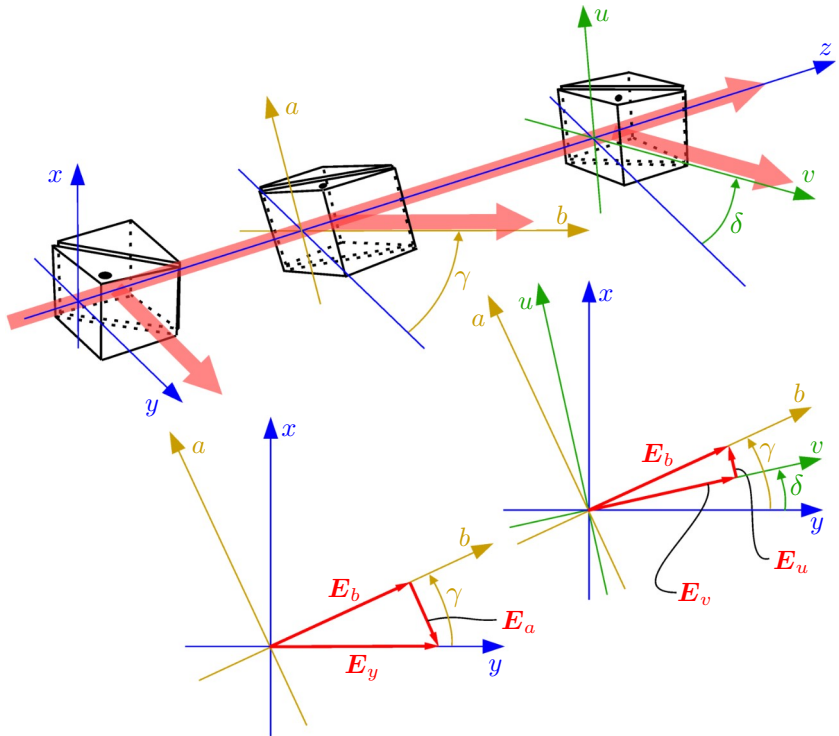


Fig. 2.9: Gedrehte polarisierende Strahlteiler

If several beam splitters, arranged in series as shown in fig. 2.9, are rotated relative to one another, then the light transmitted by the first beam splitter is partially transmitted and partially reflected by the second beam splitter. And the light transmitted by the second beam splitter is partially transmitted and partially reflected by the third beam splitter.

Quantitatively, we can see from the two right-angled triangles shown in red in fig. 2.9:¹⁴

$$\frac{|\mathbf{E}_b|}{|\mathbf{E}_y|} = |\cos(\gamma)| \quad , \quad \frac{|\mathbf{E}_a|}{|\mathbf{E}_y|} = |\sin(\gamma)| \quad (2.2a)$$

$$\frac{|\mathbf{E}_v|}{|\mathbf{E}_b|} = |\cos(\delta - \gamma)| \quad , \quad \frac{|\mathbf{E}_u|}{|\mathbf{E}_b|} = |\sin(\delta - \gamma)| \quad (2.2b)$$

The vertical bars indicate that the |modulus| is meant, i. e. the length of the red arrows representing the electric field strength; their direction is irrelevant in this equation. And when used with numbers, the modulus bars mean that negative numbers must be multiplied by -1 . Example:

$$|\cos(129^\circ - 22^\circ)| = |\cos(107^\circ)| = |-0.292| = +0.292$$

It doesn't matter whether $(\gamma - \delta)$ or $(\delta - \gamma)$ is substituted into (2.2); the result is always the same.

Rather than the ratio of field strengths, we will in the sequel much more frequently need to know what fraction of the power is transmitted by the beam splitter and what fraction is reflected. Power is the energy that flows through the beam splitter per unit of time. The power of light is proportional to the square of the field strength. Thus, we can see from (2.2):

¹⁴ $\cos(\gamma)$ denotes the “cosine of gamma”, $\sin(\gamma)$ denotes the “sine of gamma”. These functions are so frequently used in science and technology that virtually every pocket-calculator today has them implemented.

γ = polarization of incoming light

δ = angle of beam-splitter

$$\frac{\text{power of transmitted light}}{\text{power of incoming light}} = \cos^2(\gamma - \delta) \quad (2.3a)$$

$$\frac{\text{power of reflected light}}{\text{power of incoming light}} = \sin^2(\gamma - \delta) \quad (2.3b)$$

No modulus bars are needed in this equation, because the squares of the cosine and sine are always ≥ 0 . In diagram 2.10 the probabilities for transmission and reflection are shown. The sum of these two probabilities is always 1, of course, since one of the two must occur.

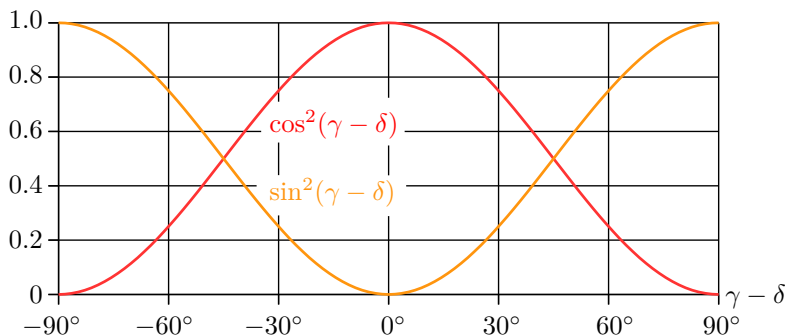


Fig. 2.10: The probabilities (2.3)

2.3 Thomas Young's Experiment

Since the 17th century, there had been two different schools of thought regarding the nature of light. One, whose most prominent proponent was Christiaan Huygens (1629–1695), favored the wave theory of light. The other, led by Isaac Newton (1643–1727),

viewed light as a stream of tiny particles. Although Huygens could present many strong arguments for his view, the dispute remained unresolved for more than a hundred years due to Newton's great authority.¹⁵

It was not until Thomas Young (1773–1829) presented his optical experiments to the Royal Society in London on November 24, 1803 [6], that the wave theory gained acceptance.

Young was an extraordinarily versatile researcher. He practiced as an ophthalmologist in London, formulated the three-color-theory of human vision, made important contributions to the decipherment of Egyptian hieroglyphs, and investigated the wave nature of sound and light.

Young persuaded his contemporaries of the wave nature of light by demonstrating in experiments that the sum of light plus light can result in darkness. This would be completely impossible if light were a stream of particles; but as the interference of waves, the observation can be easily explained.

One of his experiments, known as the double-slit experiment, later played an important role in the study of quantum phenomena in the 20th century. This is why it must be described here. It is illustrated in figure 2.11 on the next page.

A light wave, coming from the left, strikes a piece of black cardboard in which Young had pierced with a needle two small holes. In modern quantum experiments, narrow slits are used

¹⁵ Newton had a strong argument for his particle hypothesis. Waves require a medium: water waves exist only where there is water, sound waves exist only where there is air, and the rope wave in fig. 2.5 can exist only where there is a rope. In contrast, light waves propagate unimpeded even where there is no medium, i. e. in vacuum, and do so even more effectively than in transparent glass or crystals. This is easily explained if light is a stream of particles. When physicists opted for the wave theory of light at the beginning of the 19th century, they were able to explain Young's observations, but at the same time, it became impossible to understand why light can propagate in a vacuum.

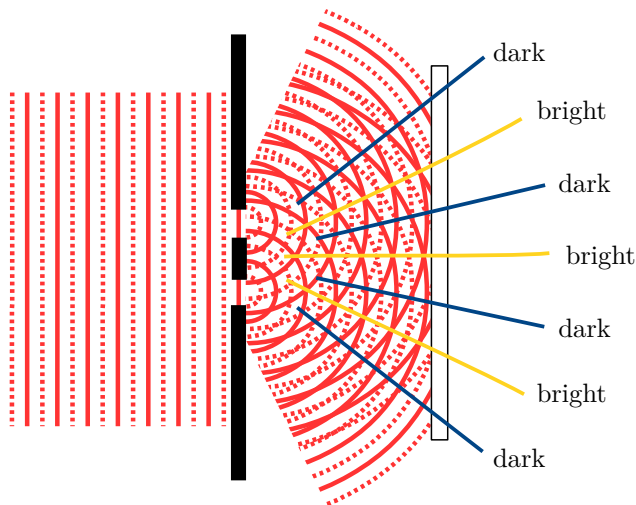


Fig. 2.11 : Thomas Young's double-slit experiment

instead of holes, hence the name “double-slit experiment”.

The solid red lines represent the crests of the light wave, while the dotted lines represent the troughs. Behind each hole in the black cardboard, a wave spreads out in all directions and interferes with the wave coming from the other hole. Yellow lines mark the areas where — as in 2.4(a) — wave crests constructively interfere with wave crests and wave troughs with wave troughs, so that bright light falls onto the white cardboard. Dark blue lines mark the areas where — as in 2.4(b) and 2.4(c) — wave crests destructively interfere with wave troughs, resulting in darkness.

Actually the matter was somewhat more complicated, because Young did not use monochromatic light, but rather the white light of the sun, which, as is well known, can be broken down into all colors of the rainbow. And light waves of different colors have different wavelengths, see table 2.1 on page 29. So Young did

not simply see light-gray and dark-gray patterns on the white cardboard, but rather patterns — offset from one another — in all colors of the rainbow. The light and dark areas of the red light fell onto different spots on the white cardboard than the light and dark areas of the green light, and these in turn fell on different spots than the light and dark areas of the blue light.

That did not change the fact that Young's observations could only be explained by the wave theory of light. The particle theory of light was thus considered to have been conclusively disproved. That is why it came as a big surprise when it, one-hundred years later, suddenly turned up again.

This page is intentionally empty.

3 Photons

3.1 The Photoelectric Effect

In 1902, Philipp Lenard (1862–1947), by then professor of physics at the University of Kiel, published the results of his investigations of the photoelectric effect [7]. As sketched in fig. 3.1, he had placed

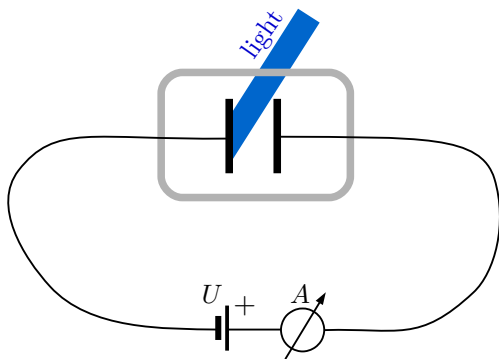


Fig. 3.1 : Measurement of the Photoelectric Effect

two metal plates parallel to each other inside an evacuated chamber, and illuminated one of them through a window in the chamber. Lenard used white light with a strong UV-component. He applied a direct current voltage U to the metal plates, and measured with an amperemeter¹⁶ the current A flowing through the vacuum vessel.

¹⁶ In fact, Lenard used an electrometer because sufficiently precise amperemeters did not yet exist at that time. This detail is irrelevant for our considerations.

How could there any current flow at all? After all, a vacuum is known to be an excellent electrical insulator. Lenard had demonstrated already earlier, however, that electrons are emitted from metal plates when light is shone on them, and then travel through the vacuum to the other metal plate. Lenard noticed:

- * Even with applied voltage $U=0$, the current is not zero.
- * The current can be reduced to zero by applying a negative external voltage. At $U \lesssim -2\text{ V}$, no electrons flow from the illuminated plate to the not illuminated plate.

Taken together, these two observations apparently imply that the electrons have a kinetic energy of up to 2 eV when they are ejected from the metal plate due to the absorbed light.¹⁷ This is because a counter-voltage of about 2 V is required to slow down the electrons and redirect them back toward the illuminated metal plate before they reach the not illuminated plate. Furthermore Lenard noted:

- * When a positive external voltage is applied, the current initially increases but then remains constant in the range from $U = 100\text{ V}$ to $U = 40\text{ kV}$.¹⁸

This, too, was easy to explain: The light releases a certain number of electrons from the metal, which, at $U = 0$, diffuse randomly through the vacuum chamber and only occasionally reach the other metal plate by chance. When a high positive voltage is applied, virtually all of the electrons are accelerated toward the not illuminated plate and collected there. No matter how much the voltage is increased, there can no more electrons be

¹⁷ Apparently, electrons can be accelerated not only by an electric voltage but also by the energy of light.

¹⁸ $1\text{ kV} = 1\text{ kilovolt} = 1000\text{ Volt}$

collected at the not illuminated plate than have been released by the light from the illuminated plate. Therefore, further increasing the voltage does not increase the current. Also this observation is plausible and easy to understand:

- * The current, i. e. the number of electrons emitted from the metal plate per unit of time, is **proportional** to the intensity of the light.

The amount of energy that the light transfers to the metal plate per unit of time is called “intensity” or “power” of the light.¹⁹ A certain amount of energy is required to emit an electron from the metal plate. The more energy per unit of time is provided by the incident light, the more electrons can leave the metal per unit of time.

There were some details, however, that were completely baffling:

- * Even with low-intensity light conditions, the current starts to flow immediately when the light is switched on, without any delay.
- * Lenard used white light with a strong UV component. When he placed between the light source and the metal plate a glass filter, which absorbed the ultraviolet portion of the light and allowed only the visible portion to pass through, then the current disappeared completely, even with arbitrarily high intensity of the visible light and with arbitrarily high external voltage U .
- * The maximum kinetic energy of about 2 eV with which the electrons are emitted from the metal is independent of the intensity of the light.

¹⁹ In this book we will use the notions “intensity” and “power” synonymously.

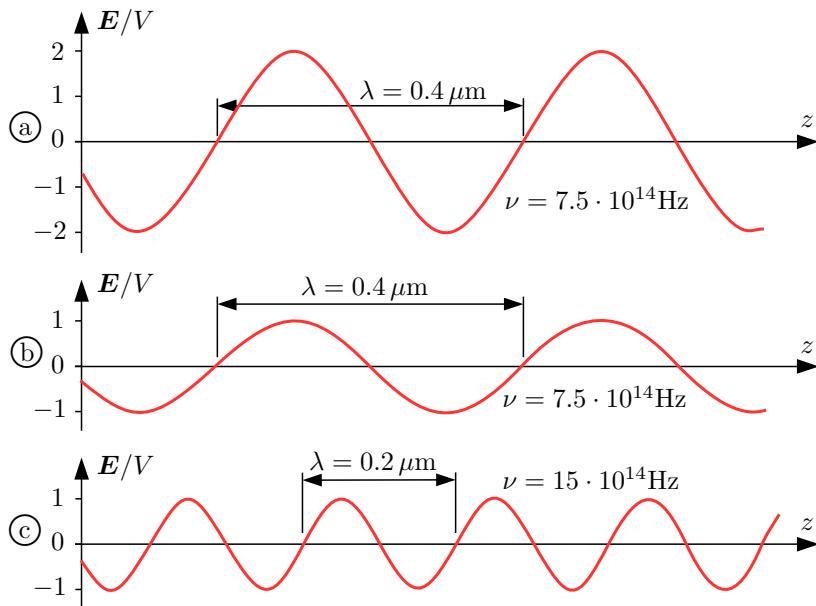


Fig. 3.2: Light-Waves

Figure 3.2 illustrates why these observations are so surprising. It shows the electric field strength \mathbf{E} of the light wave striking the metal plate. The wavelength λ of the waves 3.2(a) and 3.2(b) is $0.4 \mu\text{m}$ (violet), the wavelength of 3.2(c) is $0.2 \mu\text{m}$ (near-UV). Also plotted in the diagrams is the frequency $\nu = c/\lambda$, where c is the speed of light.

The amplitude of wave 3.2(a) is 2 V, the amplitudes of waves 3.2(b) and 3.2(c) are 1 V.

The intensity¹⁹ of a light wave is the amount of energy, which the light is transporting per time to the metal plate. The intensity is **proportional** to the square of the wave's amplitude:

$$(\text{intensity of the lightwave}) \sim (\text{amplitude of the lightwave})^2 \quad (3.1)$$

The symbol \sim means “is proportional to”, and the superscript 2 means “squared”. Thereby the ratio of the powers of the three light waves can be easily computed:

$$\frac{(\text{intensity of } 3.2\text{(a)})}{(\text{intensity of } 3.2\text{(b)})} = \frac{(\text{intensity of } 3.2\text{(a)})}{(\text{intensity of } 3.2\text{(c)})} = \frac{(2\text{ V})^2}{(1\text{ V})^2} = 4 \quad (3.2a)$$

$$\frac{(\text{intensity of } 3.2\text{(b)})}{(\text{intensity of } 3.2\text{(c)})} = \frac{(1\text{ V})^2}{(1\text{ V})^2} = 1 \quad (3.2b)$$

The wave 3.2(a) transports per time 4 times as much energy to the metal plate as the wave 3.2(c). Nonetheless electrons are emitted from the metal plate due to irradiation with wave 3.2(c), while no electrons are emitted with irradiation by wave 3.2(a).

One would actually expect that the more light energy available, the easier it would be to emit an electron from the metal. As stated in (3.1), the intensity of a wave depends solely on its amplitude; frequency plays no role. Lenard’s experiment, however, yielded this result: The frequency of the light wave is the crucial parameter; electrons are only released from the metal plate at sufficiently high frequencies. And one would actually also expect that, at low light intensity, the electron would first have to accumulate energy for some time before it can leave the metal. But in fact, the current starts immediately, even at arbitrarily low light intensity, as long as the frequency of the light is high enough.

3.2 The Hypothesis of Light-Quanta

A surprising proposal for the explanation of these puzzling observations came in 1905 from Albert Einstein (1879–1955). At that time, Einstein was employed as “technical expert of 3. class” at the Swiss Patent Office in Bern. Apparently, this job left him

enough time and energy to pursue scientific questions on the side. In 1905 he published three extraordinarily important articles: In March 1905, the “hypothesis of light-quanta” [8], which explained Lenard’s observations regarding the photoelectric effect, and for which Einstein was later awarded the Nobel Prize; in May 1905, the proof that Brownian motion could be used to experimentally verify whether atoms actually exist [9]; and finally, in June of the same year, his Special Theory of Relativity [10].

The explanation Einstein proposed for Lenard’s results was as simple as it was surprising: it boiled down to the fact that the wave model does not fully and accurately describe the properties of light. Of course, Einstein was familiar with the interference experiments that show that light plus light can result in darkness. And he knew that these experiments clearly prove that light must have the properties of waves. But at the same time, it was clear to him that the amount of energy a wave carries per unit of time depends exclusively on the wave’s amplitude, and in no way on its frequency. In the introduction to his article [8], Einstein wrote:

“The [...] undulation theory²⁰ of light has proven itself excellently in explaining purely optical phenomena, and will likely never be replaced by another theory. It must be borne in mind, however, that optical observations refer to time-averaged values, not to instantaneous values, and despite the complete confirmation of the theory of diffraction, reflection, refraction, dispersion, etc., it is conceivable that the [...] wave theory] of light may lead to contradictions with experience when applied to the phenomena of light generation and light conversion.

It now seems to me, in fact, that observations concerning ‘blackbody radiation’, photoluminescence, the production

²⁰ undulation theory = wave theory

of cathode rays²¹ by ultraviolet light, and other phenomena concerning the generation or transformation of light, appear more understandable under the assumption that the energy of light is distributed discontinuously in space. According to the assumption to be considered here, [...the energy of a light ray] is not distributed continuously [...], but consists of a finite number of energy quanta localized at points in space, which move without dividing and can only be absorbed and generated as a whole.”

So this is Einstein’s hypothesis of light-quanta: as long as we are dealing with purely optical phenomena, we should continue to think of light as a wave; but as soon as we are dealing with the production or absorption of light, we should think of light as a stream of energy quanta that “can be absorbed and produced only as a whole.”

Einstein was well aware that the concepts of waves and particles are absolutely incompatible. Light plus light can only result in darkness because the electric field strength of a wave can be positive or negative, see fig. 2.4 on page 26. In contrast, the energy of a photon is always positive. If two light quanta reach the same detector simultaneously, the detector will register twice as much energy as it would for a single light quantum. Two light quanta can never cancel each other out. But Einstein recognized more clearly than most of his contemporaries that it was hopeless to try to interpret the experimental facts without contradiction within the framework of established physical theories.

In his article [8], Einstein concerned himself on 11 pages with blackbody radiation, on just under 1 page with photoluminescence, and on 3 pages with Lenard’s investigations of electron emission due to ultraviolet light. Blackbody radiation was thus by far his most important topic. Blackbody radiation refers to the (infrared,

²¹ cathode rays = electron rays. Here Einstein refers to Lenard’s investigations.

visible, and ultraviolet) light inside a furnace.²² Its spectrum (the distribution of energy across the various frequencies of light waves) is determined by the temperature of the oven, and by nothing else. At relatively low temperature, no radiation is visible, but one can feel the infrared radiation on the skin. At higher temperature, the oven begins to glow red, and at even higher temperatures, it becomes white-hot with a high proportion of UV radiation. The theory of blackbody radiation is a fascinating topic for physicists, but far too difficult for this book.

In 1900, Max Planck (1858–1947) had discovered — as a “lucky guessed interpolation formula” [11], as he candidly admitted when receiving the Nobel Prize — the following relationship between the temperature of a furnace and the spectrum of blackbody radiation:

$$\text{energy density of black-body radiation} = \frac{8\pi h\nu^3/c^3}{e^{\frac{h\nu}{kT}} - 1} \quad (3.3)$$

No reader should get frightened by this complicated formula. We can forget about it right away; it is included here only to display the factor $h\nu$, which appears in the [exponent](#) of the number e .

h is a physical constant discovered by Planck, known as the Planck constant.²³ The formula (3.3) describes an unambiguous relationship between the temperature T of the furnace and the

²² Einstein used the term “blackbody radiation”, although it is imprecise. What is meant is: the electromagnetic radiation inside a black furnace. The walls of the furnace must be black, i. e. they must absorb and emit electromagnetic radiation of any frequency. If the walls of the oven were not black but reflective, then no thermodynamic equilibrium would be established between the radiation and the walls of the oven, and Planck’s formula (3.3) would not be valid.

²³ This is the value of the Planck constant:

$$h = 6.63 \cdot 10^{-34} \frac{\text{kg m}^2}{\text{s}} = 4.14 \cdot 10^{-15} \text{eV s} \quad (3.4)$$

electromagnetic radiation it contains with frequency ν . All other factors in this formula are constants.

Einstein devoted 11 pages to this formula to show that his light-quanta hypothesis is consistent with Planck's formula precisely if each light quantum has exactly the energy²⁴

$$E = h\nu . \tag{3.5}$$

This is the crucial point: The power¹⁹ (i. e. the energy transported per unit time) of a wave depends solely on the amplitude of the wave and has nothing to do with its frequency, see (3.1). But the energy $h\nu$ of Einstein's light quanta depends on the frequency ν of the light. This provided a way to explain Lenard's observations. Einstein reasoned as follows:

An electron in the metal plate can absorb the energy of a photon completely or partially. It is extremely unlikely, however, that an electron will absorb the energy of two (or more) photons at the same time. Therefore, when the metal plate is illuminated with light of frequency ν , then

$$E_{\max} = h\nu$$

is the maximum energy that an electron can absorb. The energy W required to emit the electron from the metal is called the work function. The work function is a few electronvolts; its exact value depends on the type of metal. In the case

$$E_{\max} = h\nu < W$$

no electron at all can escape the metal, because the absorbed energy is less than the work function. This explains why the photoelectric

²⁴ It is common practice to use the same letter E for energy and for electric field strength, even though these are two completely different quantities.

effect only occurs when the frequency ν of the light is sufficiently high. In the case

$$E_{\max} = h\nu > W$$

the excess energy of the electrons is observed as kinetic energy, i. e. the maximum kinetic energy of the electrons emitted from the metal plate is

$$\text{maximum kinetic energy} = h\nu - W . \quad (3.6)$$

Einstein showed that his equation (3.6) agrees “by order of magnitude” with Lenard’s observations. A precise check was not possible because Lenard had worked with white light, whose frequency distribution was known only very imprecisely. It was not until eleven years later that Robert Andrews Millikan (1868–1953) published the results of precise measurements of the photoelectric effect using monochromatic UV-radiation [12], which fully confirmed Einstein’s equation (3.6).

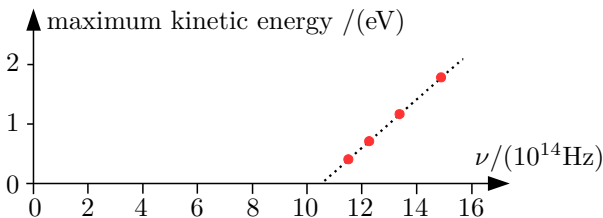


Fig. 3.3: Kinetic energy of emitted electrons

Figure 3.3 shows the results of a modern repetition²⁵ of Millikan’s experiment. A zinc plate was irradiated with four different frequencies, and at each frequency the maximum kinetic energy of

²⁵ The data come from a graph posted online by Klaus-Dieter Keller:

<https://commons.wikimedia.org/w/index.php?curid=24751457>

the electrons was measured using Lenard's counter-voltage method. The results are plotted as red points in the diagram. The dashed line drawn through the measurement points has a slope of

$$\frac{2 \text{ eV}}{4.8 \cdot 10^{14} \text{ Hz}} = 0.417 \cdot 10^{-14} \text{ eV s} \stackrel{(3.4)}{\approx} h ,$$

which — within the limits of measurement accuracy — is identical to Planck's constant. At $10.4 \cdot 10^{14} \text{ Hz}$ the dashed line intersects the axis. Accordingly, the

$$\text{work function of zinc} = h \cdot 10.4 \cdot 10^{14} \text{ Hz} = 4.3 \text{ eV} .$$

Einstein's equation (3.6) — and thus the hypothesis of light-quanta — is thus confirmed by the measurement results in fig. 3.3.

Only a quarter of a century later the name *photons* became established for Einstein's light quanta. We will use that name throughout this book.

3.3 Compton-Scattering

In the years 1922–1923, Arthur Holly Compton (1892–1962) published [13] the results of experiments that were acknowledged as strong evidence for the reality of photons. Compton directed the radiation of an X-ray tube onto a graphite sample, and examined the wavelength of the **scattered** X-rays.

Thereby he observed that the wavelength of the X-rays is greater after **scattering** than before. It is plausible that the X-rays transfer some of their energy to the graphite during scattering, and therefore have lower energy after scattering than before. But according to the wave theory of electromagnetic radiation, lower energy means a smaller amplitude. The energy loss of a wave has no effect on its wavelength.

In contrast, according to Einstein's hypothesis of light-quanta, a decrease of the energy of a photon also implies a decrease of its frequency and an increase of its wavelength:

$$\text{energy of a photon} = E \stackrel{(3.5)}{=} h\nu \stackrel{(2.1)}{=} \frac{hc}{\lambda} \quad (3.7)$$

where $\nu = \text{frequency}$, $\lambda = \text{wavelength}$

By combining the wave model and the particle model of X-rays, Compton was able to explain his observations with quantitative precision:

When a particle is scattered **elastically** off another particle that is initially at rest, as illustrated in fig. 3.4, then energy and **momentum** are conserved. This means: The sum of the energies of the two particles before the collision is equal to the sum of the energies of the two particles after the collision, and the sum of the momenta of the two particles before the collision is equal to the sum of the momenta of the two particles after the collision. To distinguish them, we mark the quantities after the collision with a prime '. We denote the particle at rest with the subscript 2, and the colliding particle with the subscript 1.

$$\text{energy conservation:} \quad E_1 + E_2 = E'_1 + E'_2 \quad (3.8a)$$

$$\text{momentum conservation:} \quad \mathbf{p}_1 = \mathbf{p}'_1 + \mathbf{p}'_2 \quad (3.8b)$$

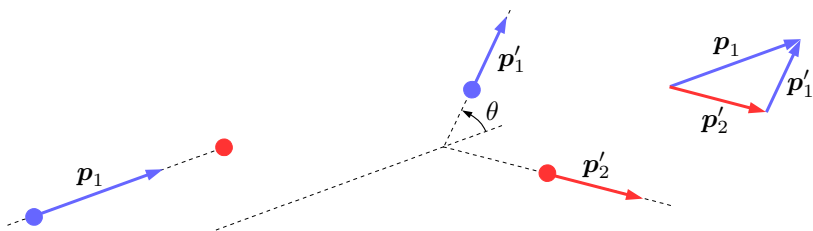


Fig. 3.4: Scattering of a particle at rest

The **momentum** \mathbf{p}_2 of the particle at rest before the collision is zero.

Regarding momentum conservation, one must consider not only the magnitudes of the momenta (represented by the lengths of the arrows) but also their directions; i. e. the arrows must be added geometrically as shown on the right in fig. 3.4. As a reminder, the momenta \mathbf{p} in (3.8) are printed in bold. When referring to the modulus (= absolute value, indicated by the lengths of the arrows), we use a regular p or vertical bars: $|\mathbf{p}| = p$

From the special theory of relativity [10], which Einstein published a few months after his hypothesis of light-quanta, the following relationship between the energy E and the magnitude p of a particle's momentum follows:

$$E = +\sqrt{c^2 p^2 + m_0^2 c^4} \quad (3.9a)$$

In this equation, c is the speed of light in vacuum, and m_0 is the mass of the particle at rest. (According to the theory of relativity, the mass of a particle is the larger the faster it moves.) The plus sign in front of the square root indicates that the positive square root is meant.

Photons differ from almost all²⁶ other particles in that their rest mass m_0 is zero. Thereby (3.9a) simplifies for

$$\text{photons:} \quad E \stackrel{(3.9a)}{=} cp \quad (3.9b)$$

Using (3.9), one can²⁷ derive without further difficulty from (3.8) the following equation¹⁴ for the scattering of a photon off a particle initially at rest with rest mass $m_0 \neq 0$:

²⁶ Only one other type of elementary particles with $m_0 = 0$ is known by today: gluons, which are responsible for the strong interaction between quarks in atomic nuclei.

²⁷ a useful exercise for physics students. All other readers should simply accept the result.

$$\frac{1}{p'_1} - \frac{1}{p_1} = \frac{1}{m_0 c} (1 - \cos \theta) \quad (3.10a)$$

Compton concentrated his investigations to X-ray photons with scattering angle $\theta = 90^\circ$. As $\cos 90^\circ = 0$, he got the simpler equation

$$\frac{1}{p'_1} - \frac{1}{p_1} = \frac{1}{m_0 c} \quad \text{if } \theta = 90^\circ. \quad (3.10b)$$

By this equation, the modulus p'_1 of the momentum of the scattered photon can be easily computed, if the modulus p_1 of the momentum of the incoming photon is known.

To check whether his measurements agreed with (3.10b), Compton had to convert the momentum of the X-rays in the particle picture back into the wavelength of the X-rays in the wave picture:

$$\text{photon momentum} = p \stackrel{(3.9b)}{=} \frac{E}{c} \stackrel{(3.7)}{=} \frac{h\nu}{c} \stackrel{(3.7)}{=} \frac{h}{\lambda} \quad (3.11)$$

With $p_1 = h/\lambda_0$ and $p'_1 = h/\lambda_\theta$, (3.10b) becomes

$$\lambda_\theta - \lambda_0 = \frac{h}{m_0 c} \quad \text{if } \theta = 90^\circ. \quad (3.12)$$

The result, which Compton published, is shown in fig. 3.5 on the next page. In the second-last line of the legend, our equation (3.12) can be recognized. Compton plotted the measured values of the radiation intensity before scattering as circles and connected them with a dashed line. The various peaks are characteristic of the material of the X-ray tube anode. Compton used a molybdenum anode, whose distinctive K_α -line lies at $0.708 \text{ \AA} = 0.0708 \text{ nm}$. The measured intensity of the radiation scattered at 90° is plotted as crosses, connected by a solid line.

As m_0 he inserted in (3.12) the rest mass of the electron, and thereby calculated a value for $\lambda_\theta - \lambda_0$ that is 10% larger than

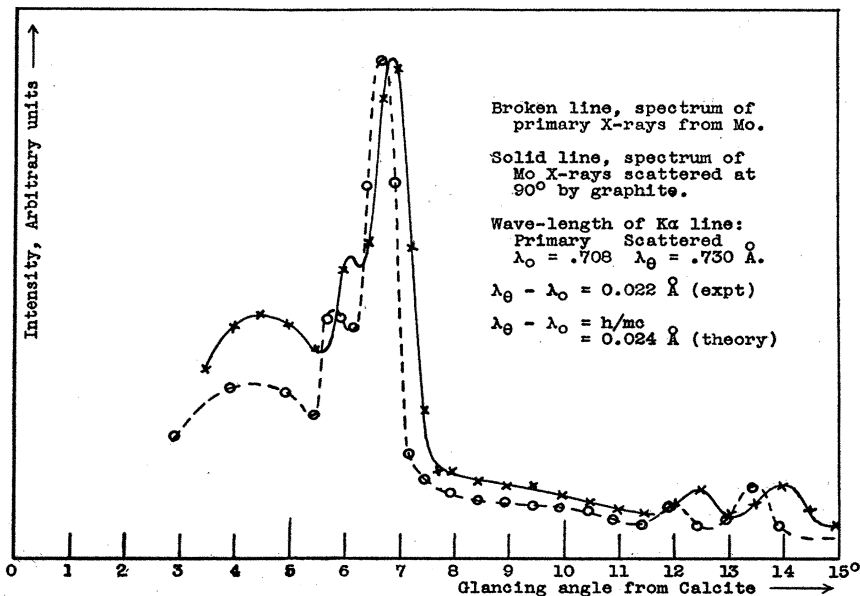


Fig. 3.5: Compton's result [13]

the measured value. The error does not result from Compton's assumption that the electrons in the graphite are at rest before the collision (which, of course, is not strictly true), nor from his ignoring the binding energy of the electrons in the graphite. These two inaccuracies are negligibly small compared to the energy of the X-ray photons.

The main error rather stems from the inaccuracy in converting the "Glancing angle from Calcite" into the wavelength of the X-rays. To determine the wavelength of the X-rays, Compton measured the angle at which they are reflected from a calcite-surface under grazing incidence. The reflection occurs due to constructive interference of the radiation scattered at various lattice planes of the crystal. To convert the scattering angle into the wavelength

of the radiation, one must know the spacing of the lattice planes in the crystal. How does one determine the spacing of a crystal's lattice planes? Through the scattering of X-rays!

This is a vicious cycle, a classic chicken-and-egg problem. To accurately measure the wavelength of X-rays, one must know the crystal geometry precisely. And to accurately measure the crystal geometry, one must know the wavelength of the X-rays precisely. In Compton's days, X-ray crystallography was still in its infancy. Given these challenges, the accuracy of his result is impressive.

3.4 Taylor's Double-Slit Experiment

In his 1905 article [8], Einstein had written that light quanta “move without splitting and can only be absorbed and produced as a whole.” Regarding the production and absorption of light quanta “only as a whole”, Einstein had presented strong arguments in support of his hypothesis. But could one also be certain that they “move without splitting”? What happens when a single photon hits a beam splitter? Is half a photon transmitted and half a photon reflected at the beam splitter, meaning it does split after all? Or will it choose one path or the other without splitting? That would mean, however, that all interference phenomena would have to disappear when experimenting with individual photons. After all, interference arises from the superposition of partial waves that have traveled different paths.

Experimenters soon began to investigate this question. Geoffrey Ingram Taylor (1886–1975) reported in 1909 [14] an interference experiment that was very similar to Young's experiment, as sketched on page 40. As light source Taylor used a gas lamp, and photographic plates instead of Young's white cardboard. First he photographed the interference pattern, then he attenuated the light using glass filters of varying darkness, and photographed the

interference pattern again. In doing so, he adjusted the exposure time so that approximately the same total amount of light reached the plate in each photograph. In the experiment with the strongest light attenuation, an exposure time of about two thousand hours (i. e. almost three months!) was required.

The reasoning was: The more the light is attenuated, the more often will it happen that only one single photon will pass through the double slit at the same time. If it does *not* split, then it will not contribute to the formation of the interference pattern, but will strike the plate completely randomly at some arbitrary point. Therefore, the interference pattern should become paler and more blurred the weaker the light intensity and, accordingly, the longer the exposure time. But that was not the case: Taylor reported that all of his images — regardless of the light intensity — showed the interference pattern with the same sharpness and clarity.

Did this prove that photons can split and thereby interfere with themselves? It took decades for physicists to understand that this proof was not really sound — and why. Light sources such as Taylor's gaslight, as well as the sun, light bulbs, or lasers, do not emit photons uniformly but in groups of varying numbers of photons, with varying intervals between the groups. It was therefore impossible to rule out with certainty that — even with extreme attenuation of the light — there were usually several photons in Taylor's apparatus at the same time, and that the attenuation of the light merely lengthened the intervals between the photon groups.

One could even go a step further and fundamentally question the concept of photons. Certainly, Einstein's hypothesis of light-quanta offered a plausible explanation for the photoelectric effect, for the spectrum of blackbody radiation, and for Compton scattering. But on the other hand, wasn't the fact that interference apparently occurs even with arbitrarily attenuated radiation an indication

that photons do not exist at all? That electromagnetic fields — as assumed in classical physics — are always continuous wave fields, and that the discontinuous and particle-like phenomena observed in the interaction of electromagnetic radiation and matter were attributable solely to some not yet understood properties of matter?

3.5 Single Photons interacting with Beamsplitters

When strongly attenuated light is observed using modern detectors, the detectors register a sequence of irregular, point-like events that are usually interpreted as photons. But could it perhaps be that what actually arrives at the detectors are not particles but wave packets, i. e. electromagnetic pulses so brief that they interact with the detectors as if they were real particles?

There is a clear distinguishing feature between — no matter how short — wave packets and photons: wave packets are split by beam splitters and are partially transmitted and partially reflected. Photons, on the other hand, are either transmitted undivided or reflected undivided. This should be easy to clarify using modern detector technology: We place detectors behind a beam splitter, allow a strongly attenuated light beam to pass through the beam splitter, and observe how often the detectors for transmitted and reflected light respond simultaneously, and how often only one of the two detectors responds.

Instead of explaining at length why it makes sense (and is even necessary) to put in a little bit more effort, I will simply describe a beamsplitter experiment conducted by J. J. Thorn et al. in 2003 [15]. As a light source, they used a method that was not developed until the 1990s and is known by the acronym SPDC = spontaneous parametric down-conversion. The method is based on the fact that in some crystals (particularly often β -barium-borate (β -Ba(BO₂)₂),

abbreviated as BBO) occasionally (“spontaneously”) converts a photon moving in a specific direction through the crystal into two photons that, taken together, have the same energy and the same momentum as the original photon. This occurs at one of approximately 10^{10} to 10^{12} photons. Thus the method is not very efficient, and a powerful laser is required to produce significant quantities of daughter photon pairs.

In most (but not all) experiments, the crystal is aligned so that each of the two daughter photons has exactly half the energy of the original photon, which is referred to as the pump photon. To conserve momentum, the pump photon and the two daughter photons must move in the same plane. This plane, however, does not need to be the same for different pairs of daughter photons. With SPDC type I, the trajectories of the daughter photons lie on a conical surface whose axis is defined by the pump beam, see fig. 3.6 on the following page. With SPDC type II, the trajectories of the daughter photons lie on different conical surfaces. Whether SPDC type I or SPDC type II occurs depends on the alignment of the crystal. In this chapter, we deal exclusively with SPDC type I.

Pinhole apertures are used in the experiments, to select pairs of daughter photons that are moving in specific directions. With SPDC type I, the two daughter photons — which we call photon_1 and photon_2 — have the same polarization, which is perpendicular to the polarization of the pump photon.

The BBO-crystal is typically 3 mm thick. Photons travel through the crystal at about half the speed they travel through air, i. e. at approximately $1.5 \cdot 10^8 \text{m/s}$. Thus the pump photon traverses the crystal within

$$\frac{3 \text{ mm}}{1.5 \cdot 10^8 \text{m/s}} = 2 \cdot 10^{-11} \text{s} .$$

Both daughter photons must be generated within this time interval. We therefore know that if the detector D (see fig. 3.6) registers a

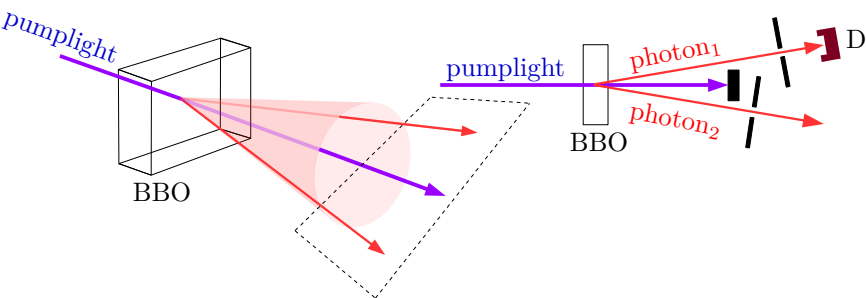


Fig. 3.6: SPDC = spontaneous parametric down conversion. With SPDC type I, the trajectories of the daughter photons (red arrows) lie on a conical surface whose axis is formed by the pump beam. The trajectories of the two simultaneously generated daughter photons and the pump beam always lie in a single plane, indicated by the dashed lines.

photon at a certain time, then the partner photon must simultaneously be located approximately at the tip of the lower red arrow in the right-hand sketch of fig. 3.6, and must be moving at the speed of light in the direction of that arrow.

If an experiment is conducted with photon_2 , one can therefore calculate exactly when and where it will be, and when it will reach a specific detector. If it is not registered by its detector at the expected time, then it has either been lost somewhere (e. g. reflected off the surface of a lens, or not reflected by a mirror), or the detector has simply missed it. This happens very frequently, because photon-detectors are quite inefficient, especially if they shall be fast. In the optical experiments presented in this book, the detectors typically have an efficiency of 10%, i. e. on average only one out of ten photons reaching the detector is actually detected.

Using SPDC type I, Thorn et al. [15] generated photon pairs to investigate exactly what actually happens at a beam splitter. The setup of their experiment is shown in fig. 3.7 on the next page. The

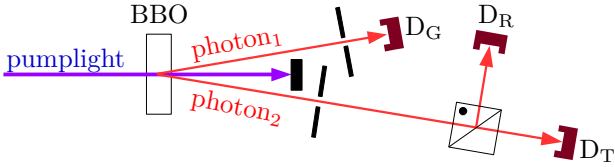


Fig. 3.7: Single photons at the beamsplitter

beam splitter was a polarizing beam splitter that was adjusted to 45° relative to the polarization of photon₂, so that the probability of transmission of the photon was equal to the probability of reflection.

Using this setup, Thorn et al. performed the following experiment exactly 100 times: For 23.4 seconds, every time when the detector D_G (the subscript G stands for gate) triggered, they checked whether within 2.5 ns the detector D_T , or the detector D_R , or both, also triggered. N_{GT} is the number of coincidences of D_G and D_T observed during these 23.4 seconds. N_{GR} is the number of coincidences of D_G and D_R . N_{GTR} is the number of coincidences of all three detectors, i. e. the number of events recorded in 23.4 seconds in which both D_T and D_R triggered within 2.5 ns after D_G had triggered. The detector D_G triggered a total of N_G times during the 23.4 seconds.

The following quantity²⁸ was calculated from the events counted within 23.4 seconds:

$$\frac{N_{GTR}/N_G}{(N_{GT}/N_G)(N_{GR}/N_G)} = \frac{N_{GTR}N_G}{N_{GT}N_{GR}} \quad (3.13)$$

This quantity may seem rather abstract at first, but it has two important advantages. First, it is independent of the efficiency of the photon detectors. For example, if the detectors have an efficiency of 10 %, then with perfect detectors (efficiency 100 %)

²⁸ For physicists: This is the $g^{(2)}(0)$ correlation. For more details, see [15].

the number N_G would be 10 times larger, and the numbers N_{GT} and N_{GR} would be 100 times larger (because a coincidence of 2 events is detected only with a probability of $10\% \cdot 10\% = 1\%$), and N_{GTR} would be 1000 times larger (because a triple-coincidence is detected only with a probability of $10\% \cdot 10\% \cdot 10\% = 0.1\%$). In (3.13), these correction factors cancel out, so there is no need to consider the efficiency of the detectors.

Second, it is proven in²⁹ appendix A.1, that (3.13) must necessarily be ≥ 1 if the energy packets registered by the detectors are short wave packets split by the beam splitter, rather than indivisible photons.

After a total of 100 runs of the experiments, each run 23.4 seconds long, the mean value and the standard deviation were computed from the results:

$$\frac{N_{GTR}/N_G}{(N_{GT}/N_G)(N_{GR}/N_G)} = \frac{N_{GTR}N_G}{N_{GT}N_{GR}} = 0.0177 \pm 0.0026 \quad (3.14)$$

The standard deviation of 0.0026 is a measure of how widely the 100 individual results are scattered around the mean value. For “normal distributed” results, the mean value of a huge number of experimental runs (the “true” value) lies with a probability of 68% less than 1 standard deviation, and with a probability of 99.7% less than 3 standard deviations from the mean value 0.0177 of the 100 conducted runs. The value ≥ 1 , which according to appendix A.1 is to be expected for wave packets, is

$$\frac{1 - 0.0177}{0.0026} \approx 378$$

standard deviations off the measured mean value. The hypothesis of divisible wave packets thus is definitively disproved by experiment.

²⁹ Here, and on several occasions in the following chapters, I move such mathematical and technical details to the appendix, which can be skipped — at least on a first reading of the book — without compromising one’s physical understanding of the subject matter.

But why are there at all triple coincidences? If light is composed of indivisible photons, shouldn't then photon₂ always arrive at only *one* detector, but never at D_T and D_R? Obviously sometimes triple coincidences happen, because (3.14) had to be zero if N_{GTR} would be zero. To clarify the issue, we estimate how often photons of two different pairs might by chance slip into the same 2.5 ns time window. Typically N_G ≈ 2.5 · 10⁶ was counted. With a detector-efficiency of 10%, we may assume that actually approximately 2.5 · 10⁷ photon pairs arrived at the detectors within the time 23.4 seconds of each experimental run. N_{GT} + N_{GR} was approximately 2 · 10⁵. Consequently, it should approximately

$$2.5 \cdot 10^7 \frac{2 \cdot 10^5 \cdot 2.5 \text{ ns}}{23.4 \text{ s}} \approx 5 \cdot 10^2$$

times happen that a “wrong” photon₂ slips by chance into the time window. With probability 1/2 it arrives at that detector which has already been triggered by the “correct” photon₂. In that case it will not be noticed. But with probability 1/2 it will reach the not yet triggered detector, and with probability 10% this detector will note the photon. Consequently about

$$5 \cdot 10^2 \cdot (1/2) \cdot 10\% \approx 25 \quad (3.15)$$

triple coincidences are to be expected due to photons which are by chance slipped into the time window. For the measured result (3.14) this implies

$$\frac{N_{GTR}N_G}{N_{GT}N_{GR}} \approx \frac{25 \cdot 2.5 \cdot 10^6}{10^5 \cdot 10^5} \approx 0.006 \quad (3.16)$$

Although this is only one-third of the measured value (3.14) = 0.0177, the order of magnitude is correct, and (3.15) is, after all, only a rough estimate. Therefore, the conclusion drawn by Thorn et al. is plausible: that (3.14) differs from zero only because of

photons that happened to slip by chance into the time window, and that the assumption of indivisible photons is consistent with this result.

3.6 Single Photons in the Interferometer

If individual photons are always either completely transmitted or completely reflected at the beam splitter, but are never split, then no interference can occur in the interferometer shown in fig. 3.8, because each photon₂ at the first beam splitter takes either path A or path B, but not both paths. Whether the photon subsequently takes path F or path H at the second beam splitter thus cannot depend on the difference in path lengths A and B.

The experiment conducted in 2004 by E. J. Galvez et al. [16] yielded a completely different result. In this experiment, correlated pairs of photons, which we call photon₁ and photon₂, were generated by SPDC type I in a BBO-crystal (β -barium-borate) approximately 6 mm thick.

Photon₁ was registered by the detector D_G. Photon₂ passed through an interferometer consisting of two standard (non-polar-

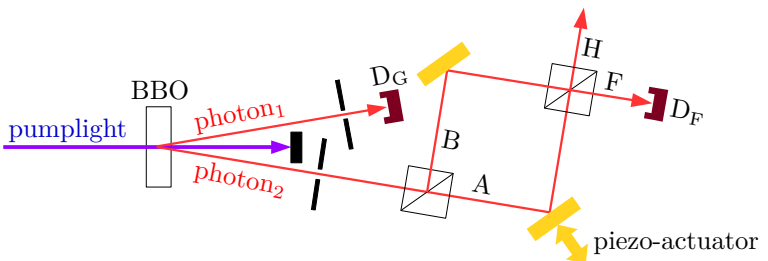


Fig. 3.8: Single photons in the interferometer

izing) beam splitters and two mirrors. One of the mirrors could be shifted by a few micrometers due to a voltage-controlled [piezo-electric actuator](#). Although the displacement of the mirror shifts the beam from this mirror to the second beam splitter by a few micrometers laterally, this is far less than 1% of the beam diameter and therefore has no noticeable effect on the interference (if any) at the second beam splitter. If the photon took path F at the second beam splitter, it could be detected by detector D_F . There was no detector at the H output. If the photon took this path, it was lost without being detected.

The experiment by Galvez et al. was conducted as follows: The voltage applied to the piezoelectric actuator, which moved one of the two mirrors, was increased in 40 equal steps from 15 V to 45 V. At each of the 41 positions of the mirror, it was counted for 20 s how often D_G and then, within 4 ns, D_F were triggered. The number N_{GF} of these coincidences is shown in the diagram in [fig.3.9](#) by red dots as a function of the piezo-voltage.

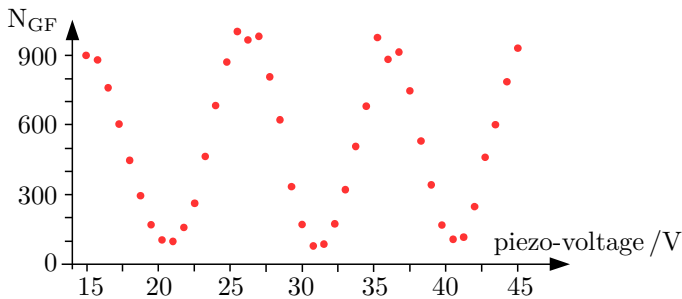


Fig. 3.9: Self-interference of single photons

The interference pattern is unmistakable. Apparently, the total length of path A was altered by just under three wavelengths, while the length of path B remained unchanged. Upon closer inspection, one notices that the patterns on the left side of the diagram are slightly more spread out than on the right side, because the

piezoelectric actuator is not as sensitive to small voltage changes at higher voltages as it is at lower voltages.

Different from diagram 2.3 on page 24, N_{GF} does not drop to 0 at the interference minima, but only to just under 100. But 2.3 was merely an idealized theoretical curve, whereas 3.9 is an unvarnished measurement result with all its experimental imperfections. Nevertheless, there can be no doubt about the reality of the interferences.

And this presents us with a problem: The diagram 3.9 shows the interference of photons that — apart from a few random coincidences — have demonstrably passed through the interferometer individually and alone. Nevertheless, all of these photons have apparently explored both paths — both path A and path B — because otherwise their behavior at the second beam splitter would not depend on the length difference between these paths. On the other hand, we saw in the previous section that individual photons always take only one path at the beam splitter, but not both paths. Is this magic, or is there a reasonable explanation?

3.7 The Position of a Particle

One is tempted to say that the photons are somehow mysteriously informed as to whether detectors are waiting for them on the other side of a beam splitter, and that they then, based on this information, either choose one of the paths, or split into both. In the following chapters we will see, however, that this view does not do justice to the facts.

Let us first take another look at the setup of the experiment by Thorn et al. in fig. 3.10 on the next page. The red arrows in the top sketch represent the trajectories of photon₁ and photon₂. A trajectory is the set of points at which a particle successively is in the course of time. We would therefore symbolize the location of a

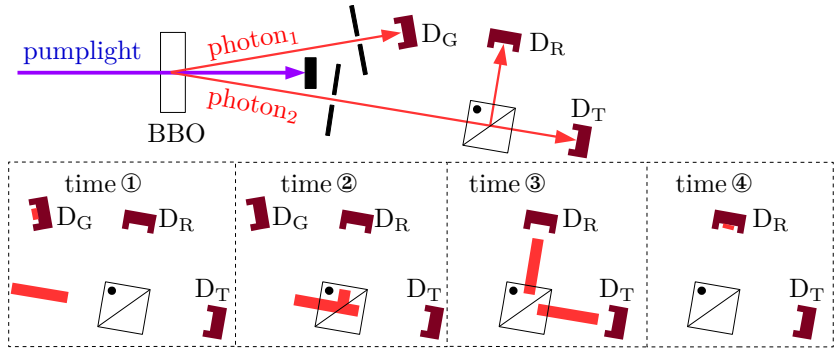


Fig. 3.10 : The location of photon₂ at different points of time

particle at a specific point in time with a red dot lying on the point of the trajectory where the particle is (presumably) located at that moment. According to classical physics, this would be correct, but according to quantum theory, that would be not correct.

In the following chapters, we will get to know compelling reasons to believe that a particle does not simply “have” a location, but rather that the location of a particle is *created* through its interaction with appropriate measuring instruments, and that the extent of the particle’s location therefore depends on the design and arrangement of the measuring instruments. The four lower sketches in fig. 3.10 are intended to illustrate this idea.

The sketch on the left, labeled time ①, shows the red painted positions of photon₁ and photon₂ at the moment the detector D_G is triggered. Photon₁ is absorbed by the active area of the detector; its position is shown in a thicker line solely to make it clearly visible. It was assumed that the active area of the detector is larger than the aperture through which photon₁ passed on its way from the BBO-crystal to the detector. Therefore, the position of photon₁ is as wide as the aperture.³⁰

At the same time the detector D_G creates by responding, in co-

operation with the SPDC-process and with the aperture through which photon_2 had to pass, also the position of photon_2 . The width of this position is determined by the aperture, its length by the time resolution of D_G . The time resolution of a photon-detector is rarely better than about 100 picoseconds. In this time, a photon travels about 3 cm, so that is the length of the position of photon_2 .³⁰

With our classically trained way of thinking, we tend to assume that the approximately 3-centimeter-long tube is the region of space within which photon_2 is located at some point — unknown to us, but in reality precisely defined. This is a misunderstanding, however, this is not what the sketch is meant to convey. Rather, according to quantum theory, the particle's location has really a surprisingly large extent. If the aperture were twice as large, then the location of photon_2 would also be twice as large. It is not the case that the photon has a specific location that is then determined with varying degrees of precision using measuring instruments of varying accuracy. Rather, according to quantum theory, the photon's location is created through the interaction between the photon and the measuring instruments, and the extent of that

³⁰ One should not take every word in this text too literally, and when looking at figures such as the “time sketches” 3.10, 4.14, or 8.14, one should “squint a bit”, i. e. not look for precision where no precision is. My aim here is to give the readers an idea of how to visualize the localizations of photons, atoms, and other quantum objects *approximately* and *in principle*, in order to be able to understand single-particle interferences. In section 9.4 I will describe an experiment, in which molecules were localized by *uncontrolled* exchange of thermal radiation with the walls of the laboratory. How should one calculate the size and shape of the location created in this way? And in section 8.3 I will introduce a “quantum eraser”, with which photons that “normally” would have caused localization, were so skillfully redirected that localization did not occur in the end. How which particular type of interaction with which specific environment creates exactly which precise shape and extent of the locations of various objects, that is — as of 2017 — not at all fully researched and not at all understood in every detail.

location depends on the design of the instruments.

The sketch “time ②” shows the position of photon₂ at the moment when it is partly already inside the beam splitter and partly still in front of it. In the sketch “time ③”, the position of photon₂ is partly still in the beam splitter, but partly already behind it. And it has split! Photon₂ has not split, but its position has.

In the sketch “time ④”, the detector D_R has triggered. As a result, the position of photon₂, which had been distributed across two paths just a moment earlier, has suddenly shrunk to the active area of D_R .

I would be *very* surprised if readers, who encounter quantum phenomena in this book for the first time, were to react to these “point in time” sketches with anything other than the utmost skepticism. It would be a great pity, however, if anyone now were to set the book aside in disappointment, dismissing it as “obvious nonsense”. For in the following chapters, I will present not only plausible reasons but also irrefutable experimental evidence that the approach to quantum phenomena suggested in these sketches is indeed appropriate to the facts. For additional motivation, I will now write down some thoughts on relational properties.

3.8 Relational Properties

Relational properties are properties that an object does not possess “by itself”, but rather in relation to other objects. Everyone immediately understands that it does not make sense to say that an object A has the property of being “larger”. A can only possess this property in relation to another object B that is smaller than A.

Unfamiliar to us is — in contrast — the idea discussed in the previous section, that a photon has a location only in relation to

the locations of other objects³¹. Many properties that an object can have “by itself” according to classical physics are treated in quantum theory as relational properties that an object can possess only in relation to its environment. The “location” of a particle is a particularly strange example, but we soon will encounter many more.

The idea will seem less strange to us once we consider, for example, whether color is a relational or an absolute property. It can happen that someone in a store carefully compares the colors of several shirts, then decides for the shirt whose color he likes best — and experiences disappointment when he wears the shirt for the first time outdoors, and notices that the shade looks quite different in sunlight than under the artificial lighting of the store, and yet again quite different under colorful disco lighting.

We tend to assume that the color we see in sunlight is the “true” color, and that color perceptions under other lighting conditions are illusions. But upon closer reflection, it quickly becomes clear that this is a rather arbitrary simplification. The objective reality is this:

The manufacturer has impregnated the shirt fabric with specific molecules that contain numerous conjugated double- and triple-bonds. These dye molecules absorb and reflect light of different wavelengths to varying degrees, and the spectrum of the reflected light then creates in the eyes of the observer a specific color impression.

Of course, the spectrum of the reflected light depends not only on the dye molecules, but also on the spectrum of the light illuminating the shirt. If the light contains a large number of photons with a specific wavelength λ_1 , and only a very small number of photons with wavelength λ_2 , then the light reflected by the shirt will contain more λ_1 photons than λ_2 photons, even if the dye molecules

³¹ which may or may not be physical measuring instruments

reflect λ_2 photons slightly more strongly than λ_1 photons. The illumination and the dye molecules together create the color of the shirt. Without dye molecules, the shirt has no color; without illumination, it has none either. Only through the interaction of dye molecules and illumination can the color of the shirt emerge, just as the location of a photon can only emerge through the interaction of the photon and suitable measuring instruments. It depends on the design and arrangement of the measuring instruments AND on the photon how extended (and possibly even split) the location of the photon is, just as it depends on the spectrum of the light AND on the chemical properties of the dye molecules which color of the shirt is produced.

The idea, that the location of an object is a relational property, is by no means a novel invention of quantum theory. On the contrary, this idea was the prevailing one for millennia, from antiquity through to the early modern period. In a tradition dating back to Aristotle (384–322), the location of an object A was not described by coordinates, but rather by “A is located between B, C, and D” or something similar.

This way of thinking was closely related to the concept of space, which people had by then. In Aristotelian philosophy, space is identical to the locations of the objects contained within it. We will easier understand that concept of space, when reading Christian Morgenstern’s (1871–1914) funny poem about the interspace:

Es war einmal ein Lattenzaun,
 Once there was a picket fence,
 mit Zwischenraum, hindurchzuschauen.
 with interspace you could see through.

Ein Architekt, der dieses sah,
 An architect who saw this,
 stand eines Abends plötzlich da –
 stood suddenly there one evening —

er nahm den Zwischenraum heraus

he took out the interspace

und baute drauß ein großes Haus.

and built a big house out of it.

Der Zaun indessen stand ganz dumm,

The fence stood there quite foolishly,

mit Latten ohne was herum.

with *nothing* in-between the pickets!

Ein Anblick grässlich und gemein.

A sight both hideous and mean.

Drum zog ihn der Senat auch ein.

Therefore the Senate confiscated it.

Der Architekt jedoch entfloh

The architect, however, fled

nach Afri- od- Ameriko. [17]

to Afri- or Americed.

It seems obvious to us that one can meaningfully speak of an interspace only in relation to the objects between which it is situated. In exactly the same way, it seemed obvious to people for millennia that one can meaningfully speak of space only in relation to the objects contained within it.

Just as we consider the interspace without the pickets, in-between which it is situated, to be absurd nonsense, so too was a space with nothing in it (i.e. a vacuum) considered absurd nonsense. That is why René Descartes (1596–1650), certainly one of the most intelligent thinkers of his time, mocked the vacuum first demonstrated in 1644 by Evangelista Torricelli (1608–1647): “If there is a vacuum anywhere, then it is in Torricelli’s head.”⁶

4 Matter-Waves

4.1 de Broglie's Hypothesis

In his light-quantum hypothesis, Einstein had proposed to consider a light wave with frequency ν and wavelength λ also as a stream of particles with energy

$$\text{photon: } E \stackrel{(3.5)}{=} h\nu \stackrel{(2.1)}{=} \frac{hc}{\lambda} \quad (4.1a)$$

and [momentum](#)

$$\text{photon: } p \stackrel{(3.9b)}{=} \frac{E}{c} \stackrel{(4.1a)}{=} \frac{h}{\lambda}. \quad (4.1b)$$

In 1924, Louis de Broglie³² (1892–1987) astonished the scientific community with the proposal to assign inversely to material particles — such as electrons — a frequency

$$\text{matter-wave: } \nu \stackrel{(4.1a)}{=} \frac{E}{h} \stackrel{(3.9a)}{=} \frac{1}{h} \sqrt{p^2 c^2 + m_0^2 c^4} \quad (4.2a)$$

and a wavelength

$$\text{matter-wave: } \lambda \stackrel{(4.1b)}{=} \frac{h}{p}. \quad (4.2b)$$

³² pronounced³³ [də brøeʒ]. Louis de Broglie was a real-life prince from an old noble family in Normandy.

³³ The IPA phonetic symbols are explained here:

<https://en.wikipedia.org/wiki/Help:IPA>

While this did restore the symmetry between waves and particles to some extent, it was wild speculation in 1924. At that time, there was not the slightest experimental evidence of wave-like properties of material particles.

Nevertheless, de Broglie's hypothesis attracted interest because many physicists hoped it would lead to a better understanding of the peculiar dual nature of photons as both waves and particles. So they considered how de Broglie's hypothesis could be tested experimentally. The hallmark of waves is interference. To induce interference in waves, one must somehow split them and then superimpose the partial waves again after they have traveled different distances.

The **momentum** of a particle at rest is $p = 0$; consequently its wavelength is infinite according to de Broglie's hypothesis (4.2b). The wavelength is finite, if the particle is moving. If an electron is accelerated by the voltage U , then its relativistic energy is

$$E = eU + m_0c^2, \quad (4.3)$$

with e being the electric charge of the electron, m_0 its rest-mass, and c the speed of light in vacuum. As these three constants are known, the energy E of an electron, which has been accelerated by the voltage U , can be computed from (4.3). Then from (4.2a) its momentum p can be computed, and thereby eventually its wavelength λ from (4.2b). The result of that computation is displayed in fig. 4.1 on the facing page.

With an accelerating voltage of less than 1 Volt, an experiment using electrons would be very difficult. When accelerated in the range from 1 V to 100 kV, electrons have the wavelength of X-rays; see table 2.1 on page 29. Therefore, it should be possible to perform interference experiments with electrons in the same way as interference experiments with X-rays, namely through **diffraction**

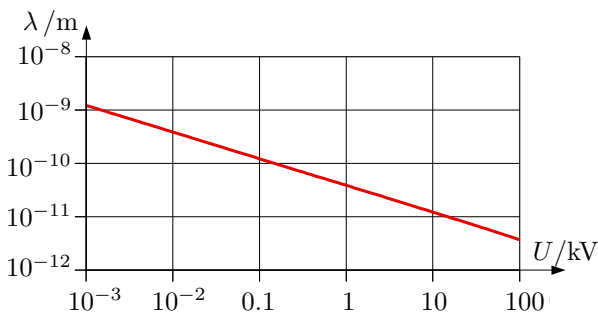


Fig. 4.1: The wavelength λ of electrons as a function of the acceleration voltage U

of the waves (assuming that electrons indeed are waves) at crystals.

4.2 Diffraction of Matterwaves by Crystals

The principle is sketched in fig. 4.2. An X-ray wave or a matter wave is traveling from top left toward a **single crystal**, which in this sketch consists of only nine atoms indicated in red. The solid blue lines indicate the wave crests, while the dashed blue lines indicate the wave troughs. The wave is scattered by the atoms and exits

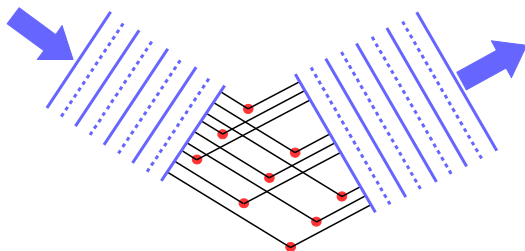


Fig. 4.2: Wave diffraction at a single-crystal

the crystal toward the upper right.

This is only possible if each of the paths, shown as black lines, is an integer multiple of the wavelength; only then do the partial waves, which take different paths through the crystal, interfere constructively.

In a real crystal, we are not dealing with 9 atoms, but with, for example, 10^{20} atoms arranged in a more or less complex three-dimensional crystal lattice. Constructive interference occurs only at a few, very specific angles. To be able to see anything at all, crystallographers often work with “white” X-rays, which contain a wide bandwidth of wavelengths.

The experimental setup is shown in the schematic diagram 4.3. Photographic film is placed in front of and behind the crystal (and often also on the sides) to record the scattered radiation. Today, instead of film usually electronic detectors are used.

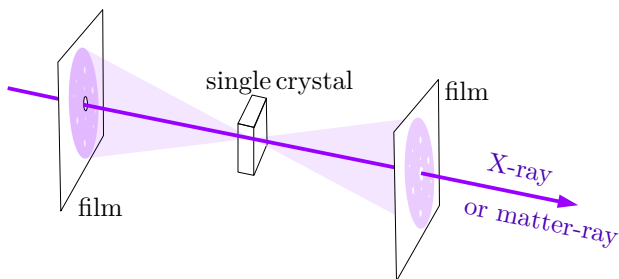


Fig. 4.3: Diffraction of X-rays by a single crystal

In fig. 4.4 on the next page, fotos³⁴ of the diffraction of electrons, X-rays, and neutrons by single crystals are displayed. These results verify de Broglie’s hypothesis (4.2) qualitatively and quantitatively.

In the image produced by X-rays, one can see the same diffrac-

³⁴ The images of the diffraction patterns were taken from websites that by today (2026) are no more online.

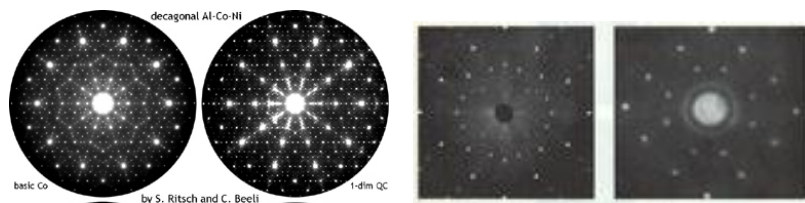


Fig. 4.4: Interference of electrons at an Al-Co-Ni alloy (2 images on the left). Interference of X-rays (3rd image) and neutrons (4th image) on identically oriented NaCl. In the center, where the matter- or X-ray beam has passed through (see fig. 4.3), there is a hole in each photograph.

tion patterns as in neutron diffraction, but also several additional patterns. Possibly a monochromatic neutron beam has been used, whereas X-ray diffraction often employs “white” radiation, which contains a wide range of different wavelengths.

Crystallographers obtain intense neutron beams from nuclear fission reactors that were built specifically for neutron research, such as the ILL in Grenoble.³⁵ It is noteworthy that deBroglie’s equations (4.2) also apply to neutrons. Unlike electrons and photons, neutrons are not elementary particles but have a complex substructure: each neutron consists of three quarks bound together by exchange of gluons.

If interference experiments can be conducted with neutrons, then they should also be possible with other composite particles, such as atoms or even molecules. This is indeed the case, but it was not achieved until the 1990s. Interference experiments with atoms and molecules are so difficult because — unlike X-ray photons, electrons, and neutrons — they cannot penetrate the lattice of crystals.

³⁵ <https://www.ill.eu/>

4.3 Neutron Diffraction by Single- and Double-Slits

One way to induce interference in larger objects is the set-up of Young's double-slit experiment, as shown in fig. 2.11 on page 40. Double slits, however, are rather crude constructions relative to the typical wavelengths of matter waves. Even with modern technology, the width of the two slits and the distance between them are significantly larger than the de Broglie-wavelength of material particles. Therefore these experiments can only be successfully carried out using highly sophisticated experimental techniques.

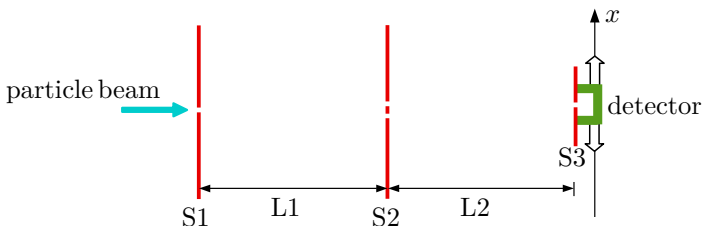


Fig. 4.5: Evaluation of interference at the double-slit

The basic setup of a double-slit experiment, which is typically located inside an evacuated chamber, is shown in fig. 4.5. The turquoise arrow represents a beam of particles entering through the entrance slit S1. After traveling a distance $L1$, the beam strikes the double-slit S2. After traveling a further distance $L2$, it strikes the entrance slit S3 of a detector that can be moved in x -direction to scan the intensity of the interference pattern.

Alternatively also detectors are used that — similar to a camera's sensor chip — can record the intensity profile of the particle beam in the detector plane with no need to move a slit. And instead of the double-slit S2, sometimes a uniform grating consisting of many parallel slits is used.³⁶

³⁶ The maxima of the interference fringes are more sharply defined with a

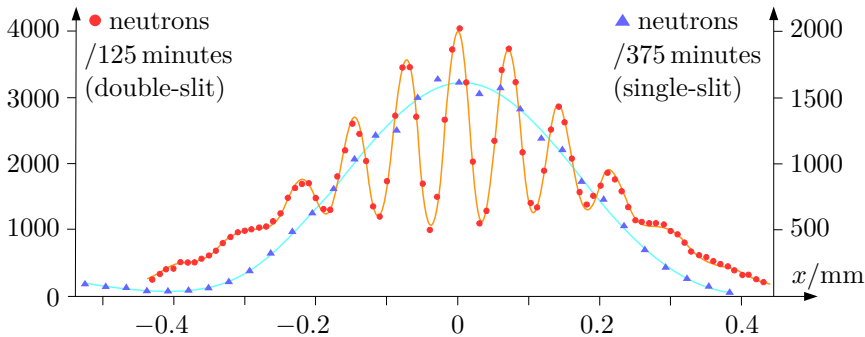


Fig. 4.6: Interference of neutrons

In 1988, Zeilinger et al. [18] demonstrated the interference of neutrons at a double slit. They used a neutron beam with a wavelength of $1.85 \text{ nm} \pm 0.14 \text{ nm}$. This wavelength was calculated from the measured velocity of the neutrons, using the de Broglie relation (4.2b). The slits S1 and S3 (see fig. 4.5) were each $20 \mu\text{m}$ wide. The double slit S2 was created by stretching a wire through the center of a $150 \mu\text{m}$ -wide single-slit. The diameter of the wire measured under the microscope was $104.1 \mu\text{m}$, the measured free width to the left of the wire was $21.9 \mu\text{m}$, and the measured free width to the right of the wire was $22.3 \mu\text{m}$. The distances L1 and L2 were 5 m each.³⁷

The red dots in graph 4.6 show the measurement results. Each individual dot represents the number of neutrons observed at this x position of the detector within 125 minutes. The solid orange line shows the expected interference pattern, calculated from the

grating of many parallel slits than with a double-slit, but their positions in the detector plane are the same as with a double-slit with same center-to-center distance.

³⁷ Anyone who hasn't yet been struck with awe should consider how difficult it is to adjust this micrometer-level precision at 5-meter distances, and to keep it adjusted over the many, many hours that the measurements took.

known geometry of the apparatus and de Broglie's equation (4.2b). One can see that the slight asymmetry of the two slit widths was taken into account in the calculation and is precisely confirmed by the measurement dots. To the right of the central maximum, four secondary maxima can be seen; on the left side, with a little generosity, even five.

The blue triangles show the result of a different measurement: For S2 (see fig. 4.5), a single slit with a width of $23\ \mu\text{m}$ was used instead of the double slit; slit S3 had a width of $60\ \mu\text{m}$, and the neutron-wavelength was $1.93\ \text{nm} \pm 0.07\ \text{nm}$. The tighter tolerance of the wavelength was achieved at the cost of the neutron flux (the number of neutrons in the beam per time) being much lower than in the double-slit-experiment.³⁸ Note that the parameters are so similar that one would have obtained almost exactly the same result if one had simply covered one of the two slits of the double-slit setup. The solid turquoise line shows the diffraction pattern calculated for this experimental setup.

The maxima and minima of the interference pattern in the double-slit experiment are absent in the single-slit experiment; thus they are clearly attributable to the fact that the neutron wave has passed through *both* slits, not just one. Even in the case of neutron-diffraction at a single slit, however, a minimum can be seen in fig. 4.6 at $-0.4\ \text{mm}$; to the left of this, a first secondary maximum begins to appear.

Zeilinger et al. [18] investigated the diffraction of neutrons at a single slit in greater detail using a $90\ \mu\text{m}$ wide single slit. In this experiment, the neutrons again had a wavelength of $1.93\ \text{nm} \pm 0.07\ \text{nm}$, and the detector slit S3 had a width of $20\ \mu\text{m}$, just as in the double-slit experiment.

³⁸ Note that the experimenters struggled for more than six hours for each single one of the small blue triangles. This measurement was an incredible strenuous effort!

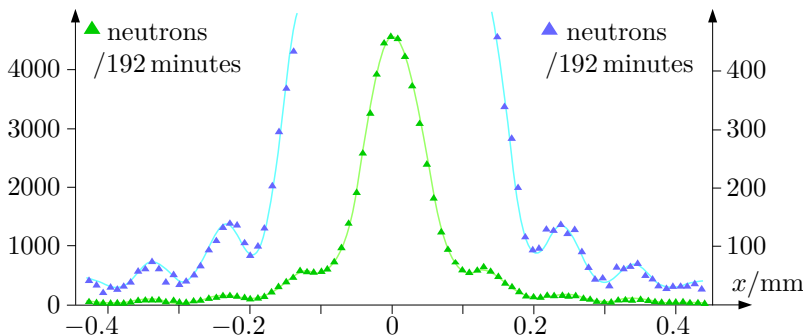


Fig. 4.7: Interference of neutrons at the single-slit

The result of this measurement is shown in fig. 4.7. Each measurement point is plotted once as a green triangle (here the left vertical axis of the diagram is valid) and once as a blue triangle (here the right vertical axis of the diagram is valid). Due to scaling the right axis by a factor of 10, the secondary maxima become more clearly visible.

Why are there at all multiple diffraction-maxima and -minima in the single-slit experiment? Doesn't in this case the entire neutron wave have to take the same path, namely through the only one available slit? How, then, can interference occur? The answer can be read from fig. 4.8 on the following page. This diagram also explains why the distance between the minima is much smaller in the double-slit experiment than in the single-slit experiment.

The sketch on the right shows two partial waves crossing the respective slit at the right edge. At an angle of β , these two partial waves are offset from each other by exactly half a wavelength. Far from the slit, these two partial waves overlap and interfere destructively, meaning they cancel each other out. Similarly, all other partial waves that are separated by a distance D cancel each other out in pairs at angle β , e. g. those that both pass through

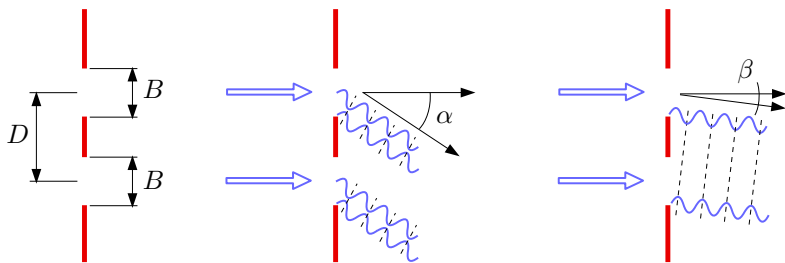


Fig. 4.8: Destructive interference at single- and double-slits

the respective slit in the middle, or both at the left edge, and so on. Overall, therefore, there is at the angle β complete destructive interference of all partial waves at the double slit.

β is the smallest angle at which total destructive interference occurs. The next instance of total destructive interference occurs when the partial waves, which are separated by a distance D , are out of phase by $3/2$ wavelengths; the next occurs when they are out of phase by $5/2$ wavelengths, and so on.

The middle sketch in fig. 4.8 shows pairs of partial waves that are separated by a distance of $B/2$. Far from the slit, they interfere with each other and cancel each other out. Similarly, all other pairs of partial waves that are separated by a distance of $B/2$ cancel each other out. As in this case the destructively interfering partial waves pass through the same slit, the interference minimum occurs at the angle α for both the double slit and the single slit. α is the smallest angle at which complete destructive interference occurs for the single slit. The next instance of complete destructive interference at the single slit occurs when the partial waves, spaced $B/2$ apart, are offset by $3/2$ wavelengths; the next occurs with an offset of $5/2$ wavelengths, and so on.

Obviously, the angles at which destructive interference occurs are inversely proportional to the distance between the interfering

partial waves:³⁹

$$\alpha \sim \frac{1}{B/2} \quad (4.4a)$$

$$\beta \sim \frac{1}{D} \quad (4.4b)$$

$$\frac{\alpha}{\beta} = \frac{D}{B/2} \quad (4.4c)$$

We can immediately check this relationship using figures 4.6 and 4.7. In diagram 4.6, the distance between the two openings of the double slit was $D \approx 126 \mu\text{m}$, and the half-width of the single slit was $B/2 \approx 12 \mu\text{m}$. Thus, according to (4.4c), $\alpha \approx \beta \cdot 126/12 \approx 10\beta$ is expected, meaning that the first minimum of the single slit should be about 10 times as far from the central maximum as the first minimum of the double slit. This is indeed the case, with 0.4 mm versus 0.04 mm. And in the experiment shown in figure 4.7, the single slit was $90 \mu\text{m}$ wide, whereas in the experiment shown in figure 4.6, it was only $25 \mu\text{m}$ wide. Therefore, according to (4.4a), the first single-slit minimum in figure 4.6 should be approximately $90/25 \approx 3.6$ times as far from the central maximum as the first minimum in figure 4.7. Again this is indeed the case, with 0.4 mm compared to $0.1 \text{ mm} \approx 0.4 \text{ mm}/3.6$.

The closely spaced minima in the red measurement curve of figure 4.6 are therefore clearly produced by neutron waves that have passed through both slits of the double slit. If a wave passes through only one slit, these minima are absent. Here we encounter the same problem we already faced with the interference of individual photons: Let's assume that the interference pattern arises because one neutron passes through the left slit, another through the right slit, and that the two neutrons then interfere in the detector plane.

³⁹ For physicists: This is an excellent approximation for small angles. Strictly speaking, the relations do not apply to α and β , but to $\sin \alpha$ and $\sin \beta$.

That certainly cannot work if one neutron follows the other at a distance of 3 meters. Even a distance of 3 centimeters would still be far too much. Let's say, with all due caution: The distance between two neutrons must in any case be much smaller than 1 cm if interference is to occur.

The speed of the neutrons was greater than 2000 m/s . Each neutron therefore traveled the distance of 1 cm in less than $5 \mu\text{s}$. Consequently, the two neutrons can only interfere with each other if they arrive at the detector within a time interval of less than $5 \mu\text{s}$. At a maximum count rate — see figure 4.6 — of approximately $4000/125 \text{ minutes} = 0.5/\text{s}$, the average time interval between two consecutive neutrons reaching the detector was 2 s. According to our assumption (maximum 1 cm separation), this average interval is 400 000 times larger than what is acceptable for interference to occur. In other words: The probability that two neutrons arrive so close together that they take different paths through the double slit and can then produce the pattern 4.6 via interference is vanishingly small. The interference pattern can only have arisen because almost all neutrons crossed the double slit *individually*, taking the path through *both* slits and interfering with themselves.

How is that supposed to work? Does half a neutron pass through the left slit and half a neutron through the right slit? How do the two halves manage to recombine into a complete neutron before they reach the detector? What happens if one half of the neutron encounters an unexpected obstacle just past the slit?

In the previous chapter, the rather surprising answer to these questions, as suggested by quantum theory, has already been presented: it is not that the neutrons “have” a location, but rather that the location of the neutrons is created by their interaction with the apparatus, particularly with the slits and detectors. The experiment by Zeilinger et al. was designed such that the position of each neutron, as it passed through the double slit, spanned both

slits, even though the neutron never splits into two particles at any time! This sounds absurd, but in the following chapters, convincing evidence will be presented to show that this is indeed the case.

4.4 Electron diffraction at single- and double-slits

In 2008, Frabboni et al. [19] reported an interference experiment with electrons, which passed through an exceptionally fine crafted double-slit. To create the double slit, the experimenters coated a 500 nm thick silicon-nitride-membrane with a 100 nm thick layer of gold. Then, using an ion-beam, two slits, each 83 nm wide and spaced (center to center) 420 nm apart, were milled into the membrane.

The electrons had the quite high energy of 200 keV, which corresponds to a speed of about $2 \cdot 10^8$ m/s (a fifth of the speed of light, after all) and a (relativistically calculated) de Broglie wavelength of 0.0025 nm. (This high energy was the reason why the SiN membrane had to be coated with gold; without this coating, the high-energy electrons would have simply shot right through the membrane even aloof the slits.) Behind the slits, the electron beam was expanded by electrostatic lenses so that the interference pattern on the detector had the same size it would have had without expansion at a distance of approximately 100 m. The detector was a 512×512 pixel CCD chip with a pixel size of $50 \mu\text{m} \times 50 \mu\text{m}$.

First, Frabboni et al. captured the top image in fig. 4.9 using an electron beam that had passed through the double slit. Then they closed⁴⁰ one of the two slits, and took the bottom image shown in fig. 4.9.

⁴⁰ It is anything but easy to close an 83 nm wide slit without affecting a slit located 420 nm away. By selectively depositing platinum onto one of the slits using an ion beam from an organometallic gas, Frabboni et al. could reliably close this slit, while the other slit only shrank from 83 nm to 76 nm.

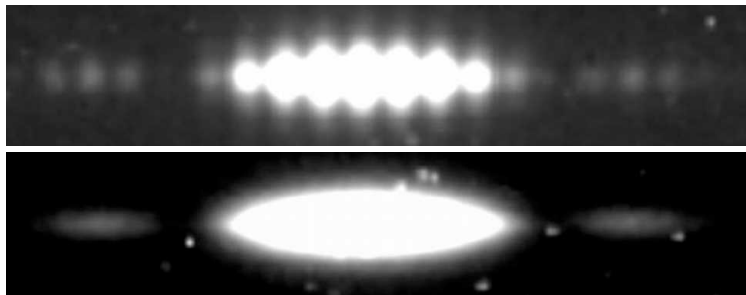


Fig. 4.9: Interference at the double-slit (top) and at the single-slit (bottom). Images from [19].

In the bottom image of fig. 4.9, one can clearly see the broad main maximum in the center, with weak secondary maxima to the left and right of it. The sub-division of these three maxima by numerous minima, which characterizes the upper figure, has disappeared in the lower figure. This result corresponds exactly to the observations made in the diffraction of neutrons at a double slit and a single slit, as shown in figure 4.7 on page 83. Using the relation

$$\frac{\alpha}{\beta} \stackrel{(4.4c)}{=} \frac{D}{B/2},$$

we can check the electron interference patterns shown in fig. 4.9 also quantitatively. The distance D between the two apertures of the double slit was 420 nm, and the width B of the single slit was 76 nm. Consequently $\alpha = \beta \cdot 420/38 \approx 10\beta$ is to be expected, i. e. the first minimum in the lower image should be about 10 times as far from the central maximum as the first minimum in the upper image. This is indeed the case. The closely spaced minima in the top image of fig. 4.9 are thus clearly produced by electron waves that have passed through both slits of the double-slit. If a wave passes through only one of the two slits, then these minima are

absent.

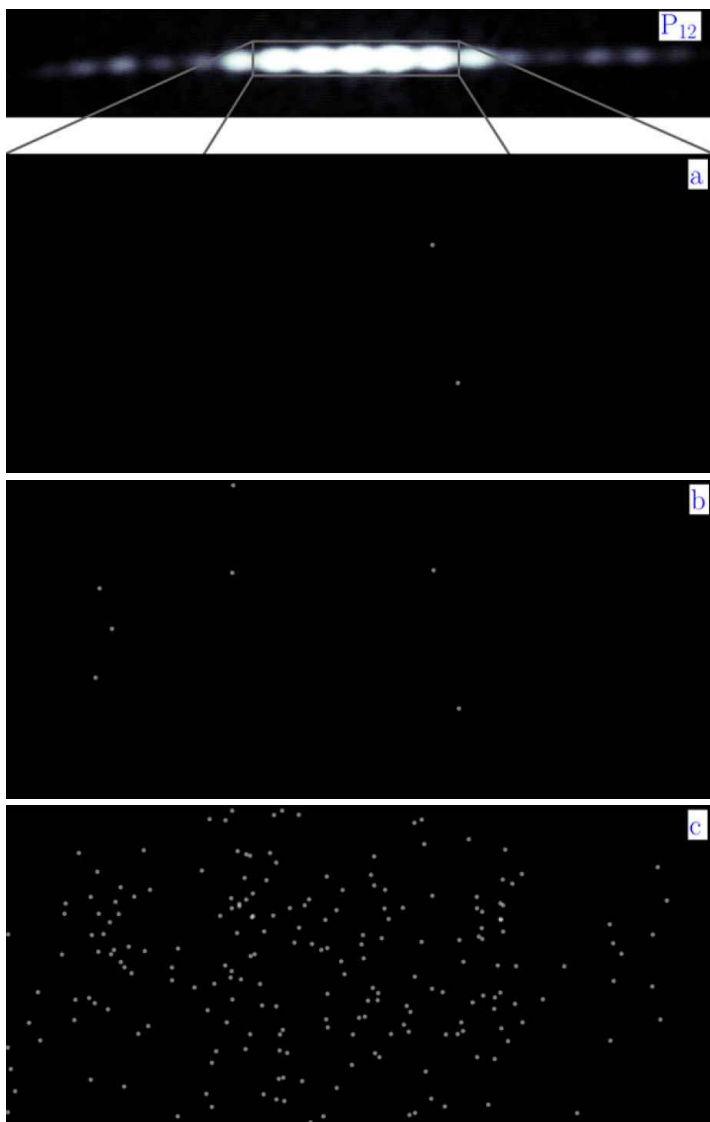
Based on the fact that each electron took about 5 ns to pass through the apparatus, while the detector registered only about once every $5 \mu\text{s}$ an electron, Frabboni et al. concluded: The probability that more than one electron was present simultaneously in the region between the electron source and the detector was negligibly small. Therefore, each individual electron must have interfered with itself. In particular, in the top image of fig. 4.9, each individual electron must have passed through both slits.

In 2013, Bach et al. [20] reported a very similar study of the self-interference of electrons at a double slit. Unlike Frabboni et al., however, they also documented the gradual emergence of the interference pattern on their CCD detector. Bach et al. accelerated the electrons with a voltage of only 600 V, resulting in the electron wavelength $\lambda = 0.0501 \text{ nm} \pm 0.0001 \text{ nm}$ according to de Broglie's relation (4.2). Each slit of their double-slit was 62 nm wide, with a center-to-center distance of 272 nm. Behind the slits, the electron beam was expanded by an electrostatic quadrupole-lens so that the interference pattern on the detector had the same size it would have had without expansion at a distance of approximately 17 m.⁴¹

The intensity of the electron beam was so low that only about 1 electron per second reached the detector. This means that the average distance between consecutive electrons was approximately $2.3 \cdot 10^6 \text{ m}$. Thus, there were practically never two (or even more) electrons simultaneously in the approximately 1 m long apparatus that could have interfered with one another. Even if, against all odds, two electrons happened to be near the double slit at the same time on rare occasions, these exotic events are statistically insignificant.

The points at which the electrons struck the detector were time-resolved recorded. The result is shown in fig. 4.10 on page 90.

⁴¹ This magnification can be read from fig. S2 on page 4 in [21].



continued on next page

continued from previous page

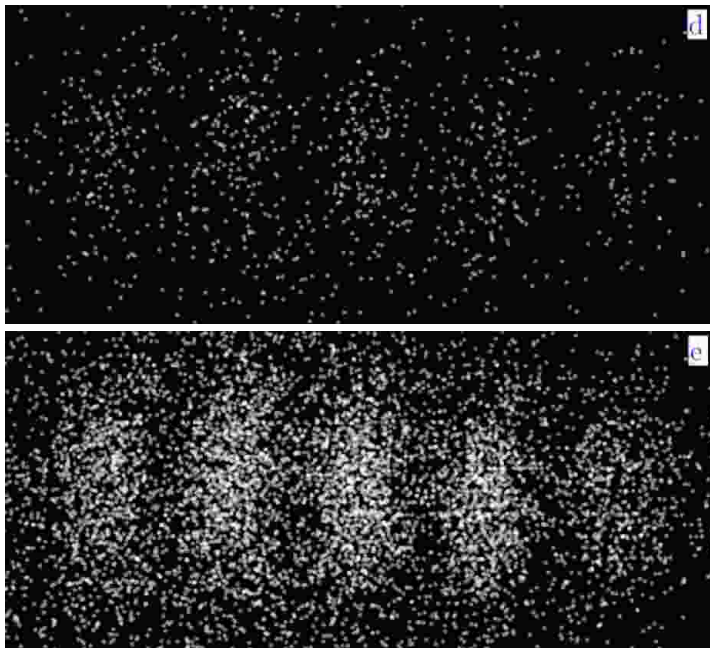


Fig. 4.10: Gradual formation of the interference pattern: The detector has been struck by (a) 2 electrons, (b) 7 electrons, (c) 209 electrons, (d) 1004 electrons, (e) 6235 electrons. This is figure 3 from [20]. The authors emphasize: “the electron detection rate in the pattern was about 1 Hz. At this rate and kinetic energy, the average distance between consecutive electrons was $2.3 \cdot 10^6$ m. This ensures that only one electron is present in the 1 m long system at any one time, thus eliminating electron-electron interactions.”

At the very top, the region of the interference pattern that was observed is indicated. This is followed by images in which the points of impact of the first 2, 7, 209, 1004, and 6235 electrons are visible as white dots. To the right and left of the central maximum, there are 3 minima on each side, which are clearly caused by destructive interference of partial waves that have passed through *different* openings of the double slit. The experiment proves that every (or at least almost every) electron passed through *both* slits individually and then interfered with itself.

4.5 Diffraction of Atoms and Molecules at the Double-Slit

The smaller the wavelength of the matter wave is relative to the width and spacing of the slits, the more difficult it is to observe interference. If one wishes to observe interference of atoms, then according to de Broglie's equation (4.2b) the best chances are with very small **momentum** of the atoms, i. e. atoms that are as light as possible and that should move very slowly.

In 1991, Carnal and Mlynek [22] observed the interference of helium-atoms flying at a speed of $990 \text{ m/s} \pm 60 \text{ m/s}$ through a double slit. The He atom is the second-lightest of all atoms. Nevertheless, at this speed, it has a wavelength of only $0.103 \text{ nm} \pm 6\%$. The other parameters — see fig. 4.5 on page 80 — of the experiment by Carnal and Mlynek were:

$$\begin{aligned}
 S1 &= 2 \mu\text{m} & L1 &= 64 \text{ cm} \\
 S2 &: \text{two slits of } 1 \mu\text{m} \text{ width each,} \\
 &\quad \text{distance (center to center)} = 8 \mu\text{m} \\
 L2 &= 64 \text{ cm} & S3 &= 2 \mu\text{m}
 \end{aligned}$$

The number of atoms detected within 10 minutes respectively at various positions of the detector along the x -axis is displayed by red

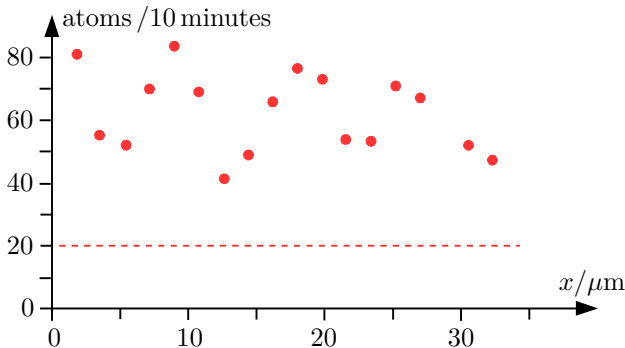


Fig. 4.11 : Interference of Helium-atoms at the double-slit

dots in diagram 4.11. The dashed line indicates how many atoms were typically counted when the entrance slit of the apparatus was closed. In an ideal experiment, the dashed line would be at zero. The “noise” of about 20 recorded atoms per 10 minutes gives an idea of the inevitable inaccuracy of this experiment.

At a wavelength of the atoms of 0.103 nm , a spacing of $8.2 \mu\text{m}$ between the interference maxima was to be expected. From diagram 4.11, Carnal and Mlynek [22] read a spacing of $8.4 \mu\text{m} \pm 0.8 \mu\text{m}$.

Neon-atoms are five times as heavy as helium-atoms. According to de Broglie’s relation (4.2b), Ne-atoms therefore have the same wavelength as He-atoms when they move five times slower. In 1991, Shimizu et al. [23] were able to demonstrate the interference of neon atoms, even though they used a detector whose x -resolution was more than a factor of 10 worse than in the experiment by Carnal and Mlynek. Shimizu et al. were nevertheless successful because the neon atoms in their experiment had a velocity of less than 2 m/s , whereas the helium atoms in the experiment by Carnal and Mlynek had a velocity of about 1000 m/s .

The setup of the experiment by Shimizu et al. [23] is outlined in

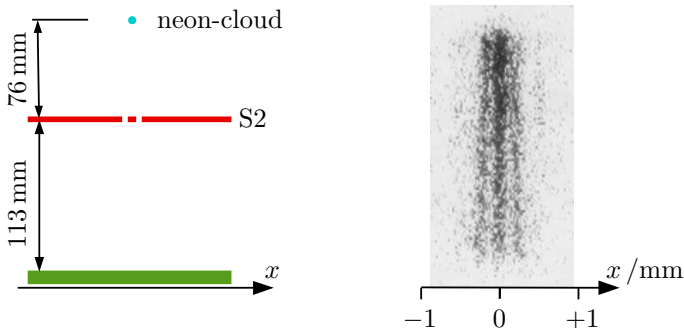


Fig. 4.12: Interference of neon-atoms at the double-slit

figure 4.12. The turquoise dot in the upper left is representing a cloud of neon atoms floating in a vacuum chamber. Under the influence of gravity, neon atoms fall 76 mm from the cloud onto a foil, into which a double-slit has been etched. Each slit is $2\ \mu\text{m}$ wide, and the distance (center-to-center) between the two slits is $6\ \mu\text{m}$. After passing through the double-slit, the atoms fall another 113 mm onto a detector, which records the point of impact with a spatial resolution of $20\ \mu\text{m} \times 20\ \mu\text{m}$ and a temporal resolution of 17 ms.

Galileo discovered already four-hundred years ago, that in a vacuum all objects fall to ground at the same speed.⁴² Under the influence of Earth's gravitation, the neon-atoms fall to ground like stones with an acceleration of $9.81\ \text{m/s}^2$. If their initial velocity is zero, then they pass through the double slit at a speed of $1.22\ \text{m/s}$ and strike the detector at a speed of $1.93\ \text{m/s}$ after a total fall time of 197 ms.

The real challenge in this experiment was to actually let the

⁴² This is remarkable because, in Galileo's lifetime, no one had yet been able to create a vacuum. Through intelligent reasoning, Galileo nevertheless discovered (and explained in the *Discorsi* [24]) that this is the case and why it cannot be otherwise.

atoms fall at a specific moment with an initial velocity of at least approximately zero. The warmer the gas, the greater the average velocity of the atoms in it. Shimizu et al. [23] used a system of four crossed laser beams, whose frequencies were carefully tuned to different absorption lines of neon, to cool the neon cloud to a temperature of 2.5 mK, which is 2.5 thousandths of a degree above absolute zero.

At a certain point in time, the lasers were switched in such a way that atoms fell out of the cloud. Ideally, they should have reached the detector 197 ms later. The right-hand image in fig. 4.12 shows, as black dots, the points of impact of atoms that reached the detector between 175 ms and 208 ms after the lasers were switched, i. e. they had already within the cloud a small initial upward or downward vertical velocity.

The interference-fringes are clearly visible. Because the interference pattern is more distinct in the lower third than in the upper part, Shimizu et al. [23] analyzed only the lower third in greater detail, as shown in diagram 4.13. The red dots indicate the total number of atoms recorded at various distances from the central interference fringe. The blue line represents the expected interference pattern as calculated by de Broglie's equations (4.2).

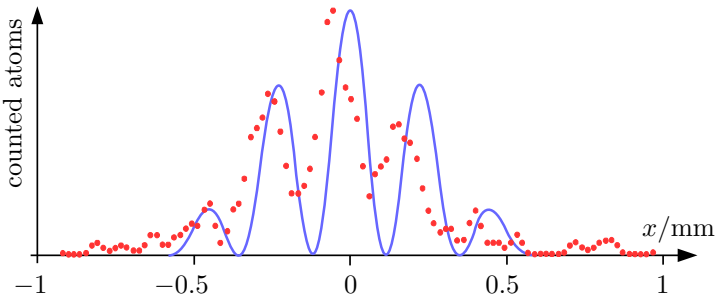


Fig. 4.13 : Interference of neon-atoms at the double-slit

The measured values exhibit a clearly discernible twist toward left. But, considering the enormous technical difficulties, they can certainly still be regarded as a convincing confirmation of de Broglie's hypothesis.

In total there are about 6 000 data points in the right image of fig. 4.12, which were obtained in more than 100 000 runs of the experiment.⁴³ On average, therefore, the experiment had to be performed more than 16 times to register just a single atom! This is yet another clear indication that it was not Ne-atoms that passed through the right slit that interfered with other Ne-atoms that passed through the left slit. Rather, obviously every — or at least almost every — atom registered by the detector must have previously crossed the double slit *individually*, using *both* slits, and must thereby have interfered with itself.

It should be noted here that the detector could only detect neon atoms that had been excited into a metastable state during laser cooling. The vast majority of atoms in the apparatus were not in this metastable state; consequently, they were neither cooled nor detected. The uncooled atoms, however, can not have contributed to the observed interference pattern because their high velocity resulted in a far too short wavelength.

Again we are faced with the puzzling phenomenon of self-interference of individual atoms: How can a single atom pass simultaneously through two slits, which are several micrometers apart? As long as we imagine (as Laplace did in the quote printed on page 12 in the first chapter) that an atom “has” a specific location (even if we do not know it exactly), this seems completely impossible.

If we acknowledge, however, that the position of the atom is determined by its interaction with the apparatus, the situation

⁴³ This, of course, requires that the experiment runs fully automatized. It was performed computer-controlled about twice per second, so that the roughly 100 000 runs took place within a total of about 14 hours.

looks quite different. Let us consider, step by step, the position³⁰ of a single Ne-atom that is detected by the detector at time ⑧, see fig. 4.14. The detector registers this atom with one of its $20\ \mu\text{m} \times 20\ \mu\text{m}$ pixels. So that is the position of this atom at time ⑧.

About 197 ms earlier, the atom began to fall from above toward the double-slit. At that moment — referred to in fig. 4.14 as time ① — the atom’s position had the shape of a sphere with about 1 mm diameter. This position was created by the atom’s interaction with the cooling lasers. Along with the position, the cooling lasers also created the initial velocity that the atom had when the lasers were turned off. This initial velocity was very small, but not zero. Therefore, the atom’s position continued to expand as it fell.

The interaction with the double-slit severely restricted the atom’s position in horizontal direction — and split it. This is the concept that we find extremely hard to grasp, one that our “common sense”, which our ancestors acquired over millions of years of evolution and passed down to us, vehemently resists.

Only the position of the atom is split; the atom itself is not

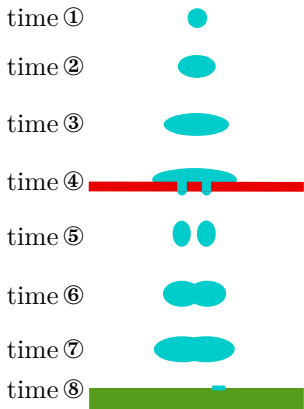


Fig. 4.14: The position of a Ne-atom at various points in time. At time ①, the atom’s position is established approximately 76 mm above the double-slit due to its interaction with the cooling lasers. Due to the interaction with the double-slit, the position of the atom at time ④ is significantly reduced, and split into two parts. Only the position of the atom is split; the atom itself remains a complete, undivided atom at all times. At time ⑧, the atom’s position shrinks abruptly to the size of a detector-pixel ($20\ \mu\text{m} \times 20\ \mu\text{m}$) due to the interaction between the atom and the detector.

split. This can be easily verified by placing the detector directly beneath the double-slit. In this case, only the pixels behind one slit are ever activated; pixels behind both slits are never activated simultaneously. Using photons as an example, this was explicitly verified in the experiments with beam splitters described in the previous chapter.

The position of the atom is jointly created by the cooling lasers, the double-slit, the detector, and the atom itself. The extent and shape of the atom's position depend on the arrangement of these devices. If the detector is placed immediately behind the double-slit, then the atom's position is always simply connected. If the detector is placed far enough away from the double-slit, then — as evidenced by the interference pattern recorded by the detector — the position of the atom is split at time ⑤ (see fig. 4.14) into two separate regions.

I will have much more to say about this in this book, and above all I will present experimental evidence that this perspective is indeed appropriate to the facts, however counterintuitive the concept of a split location of an undivided atom may seem initially. As a first indication, we can at least already observe that this perspective gives us an idea of why at all individual electrons, photons, neutrons, and atoms are capable to interfere with themselves.

Interference-experiments with still slightly heavier objects, namely sodium atoms and -molecules, were conducted in 1995 by Chapman et al. [25]. They prepared in a vacuum chamber a beam of Na-atoms and Na₂-molecules using krypton-atoms as carrier gas. All atoms and molecules in this beam had a velocity of $v = 820 \text{ m/s} \pm 3.5 \%$. Thus, according to de Broglie's equation (4.2b), the Na-atoms and Na₂-molecules had a wavelength of

$$\begin{aligned}\lambda_{\text{Na}} &= 0.021 \text{ nm} \\ \lambda_{\text{Na}_2} &= 0.011 \text{ nm} .\end{aligned}$$

The beam was directed onto a film, into which a grating³⁶ of numerous $0.03\ \mu\text{m}$ -wide slits with $0.1\ \mu\text{m}$ spacing (center to center) was etched. The incident atoms and molecules were counted 1.20 m behind the grating. The krypton-atoms were invisible to the detector. The result is plotted in the upper diagram of fig. 4.15.

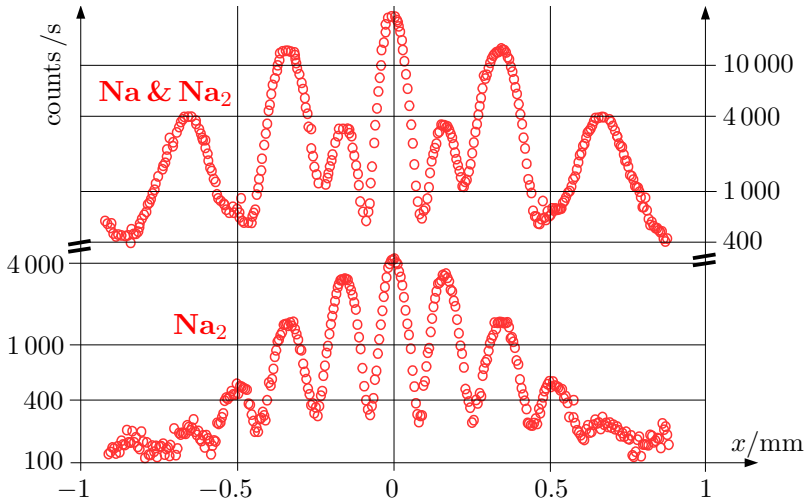


Fig. 4.15: Interference of Na & Na₂ (top) and Na₂ (bottom)

In a second experiment, the Na-atoms were before the grating scattered out of the beam by intense irradiation with a laser having a wavelength⁴⁴ of $\lambda = 0.589\ \mu\text{m}$, so that only Na₂-molecules and krypton-atoms, which do not absorb light at this wavelength, remained in the beam and passed through the grating to the detector. The result of this measurement is plotted in the bottom diagram 4.15.

As the wavelength of the Na-atoms is twice that of the Na₂-

⁴⁴ The orange line in the sodium spectrum is familiar from the lighting at crossroads. It can also be observed by sprinkling table salt (NaCl) into a candle flame.

molecules, their diffraction maxima are spread out twice as wide in the detector-plane. The

diffraction-maxima of

Na at 0 mm	± 0.33 mm	± 0.67 mm
Na ₂ at 0 mm	± 0.17 mm	± 0.33 mm
	± 0.5 mm	± 0.67 mm

are clearly visible. This experiment, too, provides striking confirmation of de Broglie's relation (4.2).

Would this also work with even heavier molecules? There is no fundamental reason why interference experiments should not be possible with objects of any mass (for example, with trucks). It is "merely" a matter of the experimenters' skill in overcoming the technical difficulties. The heaviest objects whose wave nature, according to the de Broglie-equation (4.2), could be demonstrated up to now, were molecules consisting of more than 800 atoms and weighing more than 10 000 hydrogen-atoms [26, 27].

This page is intentionally empty.

5 Vectors and Projection Amplitudes

Galileo Galilei (1564–1642), one of the founders of modern age physics in the 17th century, wrote in the sixth letter of his book “Il Saggiatore”⁴⁵ [28]:

“Philosophy [Galilei is speaking of the science of Nature] is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth.”

This remains true today, and it is especially true when we attempt to describe and understand quantum phenomena such as the puzzling self-interference of single particles. The mathematical elements we need in quantum theory to avoid “to wander about in a dark labyrinth” are not “triangles, circles, and other geometric figures”, but numbers and vectors.

Everyone knows what numbers are. As for vectors, all we need to know is that they can be represented (with some limitations, which will be explained below) by arrows, just as we have done already in earlier chapters of this book. Because working with vectors is so important, and we will encounter them so frequently in what follows, we will now examine them in a bit more detail and in a more systematic way.

⁴⁵ (Italian) *saggiatore* = gold assay balance

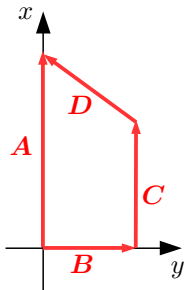


Fig. 5.1: Calculating with vectors

In diagram 5.1 one arrives, starting from the coordinate origin, on the way $\mathbf{B} + \mathbf{C} + \mathbf{D}$ at the same point as on the way \mathbf{A} . Therefore this vector-equation holds:

$$\mathbf{A} = \mathbf{B} + \mathbf{C} + \mathbf{D} \quad (5.1a)$$

If, on the other hand, we consider the length of the path, the path on the right is significantly longer than the one on the left. The length of a vector is its *modulus* or *absolute value*. So far, we have used bold letters for vectors. We denote the modulus using modulus bars or regular letters:

$$|\mathbf{A}| = A = \text{modulus (i. e. length) of the vector } \mathbf{A}$$

Thus for the moduli (i. e. the lengths) of the vectors in diagram 5.1 resp. in equation (5.1a)

$$|\mathbf{A}| = A < B + C + D = |\mathbf{B}| + |\mathbf{C}| + |\mathbf{D}| \quad (5.1b)$$

holds.

5.1 Unit Vectors

Paul Dirac (1902–1984), one of the ingenious co-founders of quantum theory, introduced the following notation for vectors:

$$|a\rangle = \mathbf{a} \quad (5.2a)$$

The bar on the left side of $|\dots\rangle$ has nothing to do with modulus bars. We will see that one quickly gets used to this convenient notation. Particularly important in quantum theory are vectors whose modulus is 1, and which therefore are called unit vectors. So that we can tell at a glance whether a vector is a unit vector or

not, I will in this book always and without exception use Dirac's notation (5.2a) for unit vectors:

$$\text{modulus of } |a\rangle = \left| |a\rangle \right| = 1 \quad (5.2b)$$

For vectors which are no unit vectors (i. e. whose modulus $\neq 1$), I will instead use the notation \mathbf{A} with bold characters:

$$\text{modulus of } \mathbf{A} = \left| \mathbf{A} \right| \neq 1 \quad (5.2c)$$

In fig. 5.2(a) two rectangular coordinate systems are indicated: One with coordinates x and y , and one with coordinates a and b . Furthermore there are in this graph four unit vectors $|a\rangle$, $|b\rangle$, $|x\rangle$, $|y\rangle$, which have the same directions as the coordinate axes with same names, and one further unit vector $|p\rangle$.

In fig. 5.2(b) the unit vector $|p\rangle$ and two vectors $|x\rangle \cdot p_x$ and $|y\rangle \cdot p_y$ are indicated. p_x and p_y are numbers. Obviously $0 < p_x < 1$, because the vector $|x\rangle \cdot p_x$ has the same direction as the unit vector $|x\rangle$, but is shorter than $|x\rangle$. And obviously $-1 < p_y < 0$, because the vector $|y\rangle \cdot p_y$ points in the opposite direction of the unit vector $|y\rangle$, and is shorter than $|y\rangle$.

In fig. 5.2(c) again the unit vector $|p\rangle$ is displayed, and in addition two vectors $|a\rangle \cdot p_a$ and $|b\rangle \cdot p_b$. Obviously $0 < p_a < 1$, because the

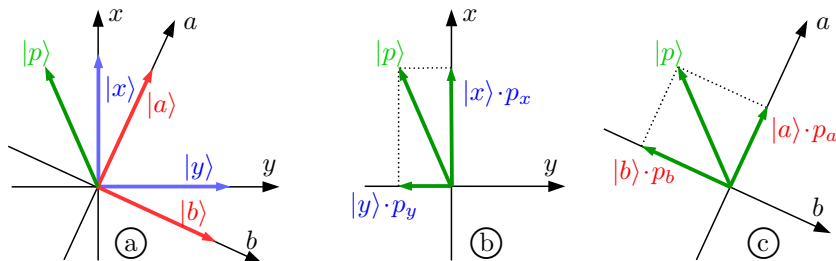


Fig. 5.2: Unit vectors and projections

vector $|a\rangle \cdot p_a$ has the same direction as the unit vector $|a\rangle$, but is shorter than $|a\rangle$. And obviously $-1 < p_b < 0$, because the vector $|b\rangle \cdot p_b$ is pointing in opposite direction of the unit vector $|b\rangle$, and is shorter than $|b\rangle$.

The numbers p_x, p_y, p_a, p_b are called *projections* or *projection amplitudes* of the vector $|p\rangle$ onto the four vectors $|x\rangle, |y\rangle, |a\rangle, |b\rangle$, respectively. For projection amplitudes, we will use the following notation:

$$p_x = \langle x|p\rangle = \text{projection amplitude of } |p\rangle \text{ onto } |x\rangle \quad (5.3a)$$

$$p_y = \langle y|p\rangle = \text{projection amplitude of } |p\rangle \text{ onto } |y\rangle \quad (5.3b)$$

$$p_a = \langle a|p\rangle = \text{projection amplitude of } |p\rangle \text{ onto } |a\rangle \quad (5.3c)$$

$$p_b = \langle b|p\rangle = \text{projection amplitude of } |p\rangle \text{ onto } |b\rangle \quad (5.3d)$$

For example, one can visualize the projection amplitude $\langle x|p\rangle$ of $|p\rangle$ onto $|x\rangle$ by illuminating the vector $|p\rangle$ from a great distance perpendicular to the x -axis in such a way that its shadow falls onto the x -axis, as indicated by the dotted line in fig. 5.2(b). The shadow arrow $|x\rangle \cdot p_x = |x\rangle \langle x|p\rangle$ is equal to the vector $|x\rangle$ multiplied by the number $\langle x|p\rangle$.

If the vector $|p\rangle$ is illuminated from a great distance perpendicular to the b -axis, then its shadow falls on the negative b -axis, as indicated by the dotted line in fig. 5.2(c). Therefore, the projection amplitude $p_b = \langle b|p\rangle$ is negative.

Vectors remain unchanged when shifted such that their magnitude and direction remain the same. Therefore, the vector $|y\rangle \langle y|p\rangle = |y\rangle \cdot p_y$ may be shifted in the coordinate plane so that it starts at the tip of $|x\rangle \langle x|p\rangle$ and ends at the tip of $|p\rangle$. And one may shift the vector $|b\rangle \langle b|p\rangle$ in the coordinate plane such that it starts at the tip of $|a\rangle \langle a|p\rangle$ and ends at the tip of $|p\rangle$, as shown in fig. 5.3.

Thus $|p\rangle$ is in each case the sum of the both other vectors:

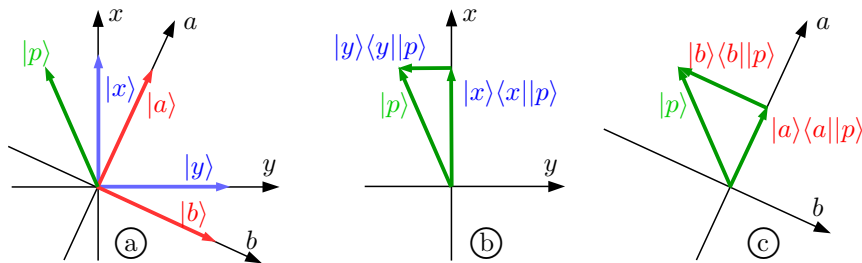


Fig. 5.3: $|p\rangle = |x\rangle\langle x||p\rangle + |y\rangle\langle y||p\rangle = |a\rangle\langle a||p\rangle + |b\rangle\langle b||p\rangle$

$$\overbrace{|p\rangle}^{\text{unit vector}} = \overbrace{|x\rangle\langle x||p\rangle}^{\text{vector}} + |y\rangle\langle y||p\rangle \quad (5.4a)$$

$$\begin{aligned} & \text{unit vector} \quad \text{projection amplitude} \\ & = |a\rangle\langle a||p\rangle + |b\rangle\langle b||p\rangle \quad (5.4b) \end{aligned}$$

Now I must mention a minor complication. Representing vectors with arrows is instructive and useful, but it is not entirely accurate. The mathematical definition of vectors is more general and comprehensive than the geometric definition of arrows. In other words: One can describe any arrow as a vector, but one cannot represent every vector by an arrow. Vectors can have mathematical properties that arrows do not have.

One distinction that is important to us concerns projection amplitudes: The projection amplitudes of arrows are positive or negative numbers, but in any case they are real numbers. In contrast, the projection amplitudes of vectors can also be complex. To clarify what complex numbers are all about, let's calculate the square root of -42.25 as an example:

$$\sqrt{-42.25} = \sqrt{(-1) \cdot (+42.25)} = \underbrace{\sqrt{-1}}_{\pm i} \cdot \underbrace{\sqrt{+42.25}}_{\pm 6.5} = \pm i 6.5 \quad (5.5a)$$

In order to be able to solve equations that contain the roots of negative numbers, mathematicians in the sixteenth century

invented the number $i = \sqrt{-1}$. The inverse of (5.5a) is

$$(+6.5) \cdot (+6.5) = +42.25 \quad (5.5b)$$

$$(-6.5) \cdot (-6.5) = +42.25 \quad (5.5c)$$

$$(+i) \cdot (+i) = -1 \quad (5.5d)$$

$$(-i) \cdot (-i) = -1 \quad (5.5e)$$

In contrast, if the signs of the two numbers are different, then

$$(+6.5) \cdot (-6.5) = -42.25 \quad (5.5f)$$

$$(+i) \cdot (-i) = +1 \quad (5.5g)$$

holds. All numbers that contain the factor $i = \sqrt{-1}$ are called *complex* numbers. All numbers that do not contain i are called *real* numbers.

$$i728.73 \quad , \quad 5.17 - i54 \quad , \quad 16 + i3.77 \quad , \quad \dots$$

are examples for complex numbers. The same rules apply to calculations with complex numbers as to calculations with real numbers. Example:

$$\begin{aligned} (5.17 - i54)^2 &= (5.17 - i54)(5.17 - i54) = \\ &= 5.17^2 \underbrace{-5.17 \cdot i54 - i54 \cdot 5.17}_{-i2 \cdot 5.17 \cdot 54} + \underbrace{i^2 54^2}_{-54^2} \end{aligned} \quad (5.5h)$$

The *complex conjugate* of a number is the number obtained by replacing all i with $-i$. The complex conjugate is denoted by a superscript asterisk $*$. Example:

$$(5.17 - i54)^* = 5.17 + i54 \quad (5.5i)$$

To compute the square of a number, the number is multiplied by itself. In contrast, to compute the *modulus square*, the number is multiplied by its complex conjugate:

$$\begin{aligned}
 & \text{modulus square of } 5.17 - i54 = \\
 & = \left| 5.17 - i54 \right|^2 = (5.17 - i54)^* \cdot (5.17 - i54) = \\
 & = (5.17 + i54) \cdot (5.17 - i54) = \\
 & = 5.17^2 - \underbrace{5.17 \cdot i54 + i54 \cdot 5.17}_0 - \underbrace{i^2 54^2}_{+54^2} = 5.17^2 + 54^2 \quad (5.5j)
 \end{aligned}$$

While the square (5.5h) of a complex number in general is again a complex number, the modulus square (5.5j) of a complex number is always a positive real number.

(5.5) contains all the rules for working with complex numbers that we need in this book. Not really difficult, but certainly needs getting used to. Every time we encounter a complex number I will therefore remind the readers of these rules.

Physicists have used complex numbers long before the discovery of quantum theory in 1925. But merely for convenience, since complex numbers allow for very elegant and efficient solutions of many mathematical problems in classical physics. If someone wanted to avoid complex numbers, however, he could still perform all physical calculations using real numbers. That changed with quantum theory. In quantum theory, complex numbers are indispensable, and that is why we cannot always avoid them in this book.

Now let's return to the projection amplitudes. *If* two vectors $|g\rangle$ and $|h\rangle$ can be represented by arrows (which is sometimes, but not always, the case), then the projection amplitude of one onto the other is equal to the projection amplitude of the other onto the one:

$$\text{sometimes } \langle g||h \rangle = \langle h||g \rangle \quad (5.6a)$$

This is often not the case, however, with the vectors we deal with in quantum theory. Rather,

$$\text{often } \langle g||h \rangle \neq \langle h||g \rangle, \text{ but} \quad (5.6b)$$

$$\text{always } \langle g||h \rangle = (\langle h||g \rangle)^* \quad (5.6c)$$

holds. The asterisk* denotes the complex conjugate, as explained in (5.5i). (5.6a) holds for real projection amplitudes, (5.6b) for complex projection amplitudes. Why (5.6c) holds cannot be explained by diagrams with arrows. We simply take note of (5.6c) as a mathematical fact.

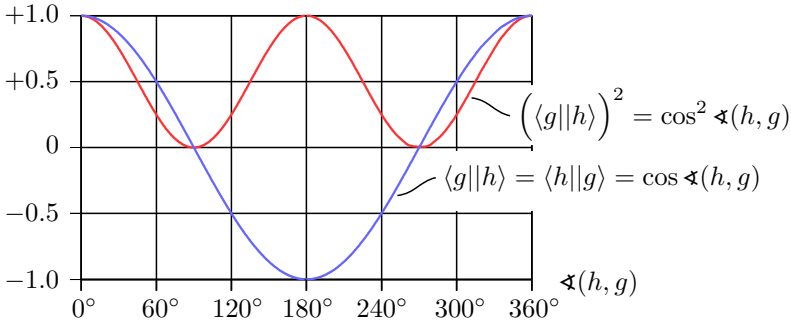


Fig. 5.4: **Real projection amplitudes** and the **square of real projection amplitudes** of unit vectors

For the special case of real projection amplitudes — i.e. the case where (5.6a) holds — the projection amplitudes $\langle h||g \rangle$ and the square $(\langle h||g \rangle)^2$ of the unit vectors $|h\rangle$ and $|g\rangle$ are plotted in diagram 5.4 as a function of the angle $\sphericalangle(h, g)$ between them: If the projection amplitudes of unit vectors are real, then they are just the cosine (abbreviated as \cos)¹⁴ of the angle between them.

if $\langle h|g \rangle$ is real:

$$\langle h|g \rangle = \langle g|h \rangle = \cos \angle(h, g) \quad (5.7a)$$

$$\left(\langle h|g \rangle\right)^2 = \left(\langle g|h \rangle\right)^2 = \cos^2 \angle(h, g) \quad (5.7b)$$

If the angle between two unit vectors is 90° or 270° , i. e. a “right angle”, then the vectors are said to be **orthogonal**. For example, $|x\rangle$ in fig. 5.3 is orthogonal to $|y\rangle$, and $|a\rangle$ is orthogonal to $|b\rangle$. The projection amplitude (i. e., the cosine) of orthogonal vectors is obviously zero: If one illuminates the vector $|x\rangle$ from a great distance perpendicular to the y -axis, then its shadow on the y -axis has length zero.

If a projection amplitude $\langle g|h \rangle$ is complex, then the vectors $|g\rangle$ and $|h\rangle$ cannot be represented by arrows, and therefore the angle between them cannot be determined. It is common practice, however, to refer to such vectors $|g\rangle$ and $|h\rangle$ as orthogonal if their projection amplitude is zero.

5.2 Two Examples

Now we have all the tools necessary for the systematic description of measurements in quantum theory. As an example, we consider the measurement of a photon’s polarization, as shown in fig. 5.5 on the next page.

The object on which the measurement is to be performed is the incident photon, which is polarized in the p -direction. The unit vector $|p\rangle$ describes the state of the photon; therefore, $|p\rangle$ is also referred to as *state vector*. The measuring device — i. e. the polarizing beam splitter with the two detectors at its outputs — is rotated such that it **reflects** x -polarized light and transmits y -polarized light.

If the photon is transmitted, it will leave the beam splitter in the state $|y\rangle$, and will be detected by detector D_T . If the photon

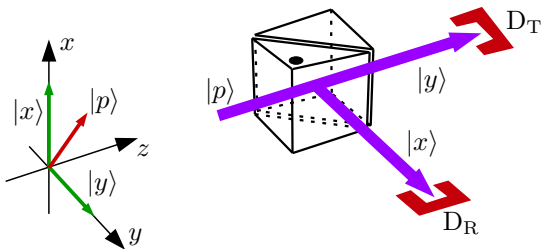


Fig. 5.5: A photon at a polarizing beam-splitter

is reflected, it will leave the beam splitter in the state $|x\rangle$, and will be detected by detector D_R . Only if $|p\rangle = |y\rangle$ or $|p\rangle = |x\rangle$ will the state vector of the photon behind the beam splitter be the same as before the beam splitter.

In classical physics, a well-conducted measurement does not alter the object being measured. One determines a property that the object already possessed before the measurement and that it still possesses after the measurement. When measuring quantum objects, the situation is different. The measuring device imposes a state on the object being measured. It is said that the measuring device prepares the object in a specific state. Every measuring device has only a limited set of state vectors that it can impose on the object. These vectors are called the *eigenvectors* of the measuring device. The polarizing beam splitter has only two eigenvectors: $|x\rangle$ and $|y\rangle$.

In which of its eigenvectors will the measuring device (the polarizing beam splitter) prepare the object (the photon)? To calculate this, quantum theory proceeds as follows: First, one writes the state vector of the object being measured (in this case, the photon) as the sum of the projections onto all eigenvectors of the measuring device:

state vector of the object before measurement =

$$\begin{aligned}
 &= |p\rangle = \underbrace{|x\rangle}_{\text{eigenvectors of the measuring device}} \overbrace{\langle x|p\rangle}^{\text{projection amplitudes}} + \underbrace{|y\rangle}_{\text{eigenvectors of the measuring device}} \overbrace{\langle y|p\rangle}^{\text{projection amplitudes}} \quad (5.8)
 \end{aligned}$$

The calculation of (5.8) has been illustrated in fig. 5.2(b) on page 104 and fig. 5.3(b) on page 106. Both possibilities are included in the state vector (5.8): reflection and transmission of the photon. But only one of the two detectors — either D_R or D_T — will detect the photon. Which one? The answer to this question is provided by Born's rule, discovered in 1926 by Max Born (1882–1970) and named after him:

Born's rule: The probability that a particular measurement result will occur is equal to the modulus square of the projection amplitude of the state vector of the measured object onto the respective eigenvector of the measuring device. (5.9)

That sounds terribly complicated, but it's actually quite simple. Let us consider the example in (5.8). The photon reaches the detector D_R if the beam splitter prepares it in the state $|x\rangle$. The projection amplitude of the photon's state vector onto this eigenvector of the measuring device is $\langle x|p\rangle$. Thus, according to Born's rule, the probability $P(D_R)$ that the photon reaches the detector D_R is equal to

$$P(D_R) = \left| \langle x|p\rangle \right|^2. \quad (5.10a)$$

In contrast, the photon will reach the detector D_T if the beam splitter prepares it in the state $|y\rangle$. The projection amplitude of the photon's state vector onto this eigenvector of the detector is $\langle y|p\rangle$. Consequently

$$P(D_T) = \left| \langle y | p \rangle \right|^2 \quad (5.10b)$$

is the probability that the photon will be transmitted to detector D_T .

If $|p\rangle = |x\rangle$, then $P(D_R) = 1$ and $P(D_T) = 0$. And if $|p\rangle = |y\rangle$, then $P(D_R) = 0$ and $P(D_T) = 1$. In all other cases, chance determines whether the photon is prepared in the state $|x\rangle$ or $|y\rangle$: We then only know that it is prepared in the state $|x\rangle$ with probability $P(D_R)$ and in the state $|y\rangle$ with probability $P(D_T)$, but we cannot predict the outcome of the measurement with certainty in any individual case. In quantum theory, measurement is the only point at which chance comes into play. As long as no measurement takes place, quantum theory is a completely deterministic theory, just like the theories of classical physics.

As a second example we consider the diffraction of neutrons at a double slit. The basic setup of the experiment is shown in fig. 4.5 on page 80, and the measurement results are represented by red dots in the diagram 4.6 on page 81. We denote the state vector of a neutron entering the apparatus through slit S1 as $|S1\rangle$. The double slit prepares it in one of its two eigenstates, $|S2_l\rangle$ or $|S2_r\rangle$. Here, the index r means that the neutron has passed through the right slit, and the index l means that the neutron has passed through the left slit. The most likely possibility is, of course, that the neutron does not hit either slit but gets stuck in the first chamber. Here we are only considering those cases, however, in which the neutron actually reaches the detector plane.

To clarify what happens to the neutron at the double slit, we describe its state (before the double slit) as the sum of its projections onto the two eigenvectors of the double slit:

$$|S1\rangle = |S2_l\rangle \langle S2_l | S1 \rangle + |S2_r\rangle \langle S2_r | S1 \rangle \quad (5.11)$$

If detectors were positioned immediately behind the two slits — as in the setup 5.5 detectors were positioned behind both outputs of

the beam splitter — then the double slit would prepare the neutron according to Born's rule (5.9) with a probability of

$$P_l = \left| \langle S_{2l} | S_1 \rangle \right|^2 \quad (5.12a)$$

in the state $|S_{2l}\rangle$, and the detector behind the left slit would register the neutron. With probability

$$P_r = \left| \langle S_{2r} | S_1 \rangle \right|^2 \quad (5.12b)$$

the double slit would prepare the neutron in state $|S_{2r}\rangle$, and the detector behind the right slit would register the neutron. Since the sum of the probabilities of all possible outcomes is always 1, and since the experiment was carefully designed so that both outcomes are equally likely,

$$P_l = P_r = 1/2 . \quad (5.12c)$$

This does *not* imply, however, that the two projection amplitudes are equal to $\sqrt{1/2}$, because projection amplitudes in general are complex numbers. Therefore we introduce the abbreviations

$$l = \langle S_{2l} | S_1 \rangle \quad , \quad r = \langle S_{2r} | S_1 \rangle \quad (5.13)$$

with $\left| l \right|^2 = \left| r \right|^2 = P_l = P_r = 1/2 .$

Thereby the state vector of the neutron can be written as

$$|S_1\rangle = l |S_{2l}\rangle + r |S_{2r}\rangle . \quad (5.14)$$

Now we have arrived at one of the strangest features of quantum theory: As there are no detectors placed immediately behind the two slits, the neutron is — according to quantum theory — *not* prepared in one of the states $|S_{2l}\rangle$ or $|S_{2r}\rangle$, but continues its flight

in the “undecided” state (5.14). The same is true for the polarizing beam splitter: If there were no detectors behind the two outputs of the beam splitter, then — according to quantum theory — the beam splitter would prepare the photon neither in state $|x\rangle$ nor in state $|y\rangle$, but rather in the “undecided” state

$$|5.8\rangle = |x\rangle\langle x||p\rangle + |y\rangle\langle y||p\rangle . \quad (5.15)$$

It seems as though what happens at the beam splitter or the double slit depends on whether or not the process is observed by detectors. The better way to look at this, however, is the one already suggested in fig. 4.14 on page 97: The position of the neutron is created by its interaction with the apparatus. The trajectory of the neutron in this experiment is the sequence of the positions it occupies one after another on its path through the slits to the detector. The state vector (5.14) reflects the fact that the trajectory of the neutron in this experiment is *not* restricted to one of the two slits, but extends over both slits.

The neutron detector scans with step size 0.01 mm the x -axis at 87 positions in total. We call the 87 eigenvectors of the detector

$$|x_1\rangle , |x_2\rangle , |x_3\rangle , \dots , |x_{86}\rangle , |x_{87}\rangle . \quad (5.16)$$

To clarify what happens to the neutron during detection, we describe its state (immediately before it strikes the detector plane) as the sum of its projections onto the detector’s 87 eigenvectors:

$$\begin{aligned} |(5.14)\rangle &= l|S2_l\rangle + r|S2_r\rangle = |x_1\rangle\left(l\langle x_1||S2_l\rangle + r\langle x_1||S2_r\rangle\right) + \\ &+ |x_2\rangle\left(l\langle x_2||S2_l\rangle + r\langle x_2||S2_r\rangle\right) + \dots + \\ &+ |x_{87}\rangle\left(l\langle x_{87}||S2_l\rangle + r\langle x_{87}||S2_r\rangle\right) = \\ &= \sum_{j=1}^{87} |x_j\rangle\left(l\langle x_j||S2_l\rangle + r\langle x_j||S2_r\rangle\right) \end{aligned} \quad (5.17)$$

In the last line, the Greek character $\Sigma = \text{Sigma}$ is used as summation symbol. The symbol is read as “the sum from j equal 1 to 87”. This means that in the first summand, the number 1 is to be substituted for j , in the second summand the number 2, in the third summand the number 3, and so on, until finally, in the last summand, the number 87 is substituted for j . The last line of (5.17) is therefore merely a shorthand notation for the lines above it.

Thereby we can calculate the probability that the detector will register the neutron at a given position. For example, the probability $P(x_{31})$ that the neutron will be observed at detector position x_{31} , is according to Born’s rule (5.9) equal to the modulus square of the projection amplitude onto the detector’s eigenvector $|x_{31}\rangle$ in equation (5.17), i. e.

$$\begin{aligned}
 P(x_{31}) &= \left| l \langle x_{31} | S_{2l} \rangle + r \langle x_{31} | S_{2r} \rangle \right|^2 \stackrel{(5.5j)}{=} \\
 &= \left(l \langle x_{31} | S_{2l} \rangle + r \langle x_{31} | S_{2r} \rangle \right)^* \left(l \langle x_{31} | S_{2l} \rangle + r \langle x_{31} | S_{2r} \rangle \right) \stackrel{(5.6c)}{=} \\
 &= \left(l^* \langle S_{2l} | x_{31} \rangle + r^* \langle S_{2r} | x_{31} \rangle \right) \left(l \langle x_{31} | S_{2l} \rangle + r \langle x_{31} | S_{2r} \rangle \right) \stackrel{(5.13)}{=} \\
 &= \underbrace{l^* l \langle S_{2l} | x_{31} \rangle \langle x_{31} | S_{2l} \rangle + l^* r \langle S_{2l} | x_{31} \rangle \langle x_{31} | S_{2r} \rangle +}_{\frac{1}{2} |\langle x_{31} | S_{2l} \rangle|^2} \\
 &\quad + r^* l \langle S_{2r} | x_{31} \rangle \langle x_{31} | S_{2l} \rangle + \underbrace{r^* r \langle S_{2r} | x_{31} \rangle \langle x_{31} | S_{2r} \rangle}_{\frac{1}{2} |\langle x_{31} | S_{2r} \rangle|^2} . \quad (5.18)
 \end{aligned}$$

In fig. 5.6 on the next page, the probability $P(x) = (5.18)$ of detecting the neutron at an arbitrary position x is plotted as the red curve; the contributions of the various terms in (5.18) are also shown in blue and green. In calculating these curves, it was assumed that the distance between the two slits (center to center)

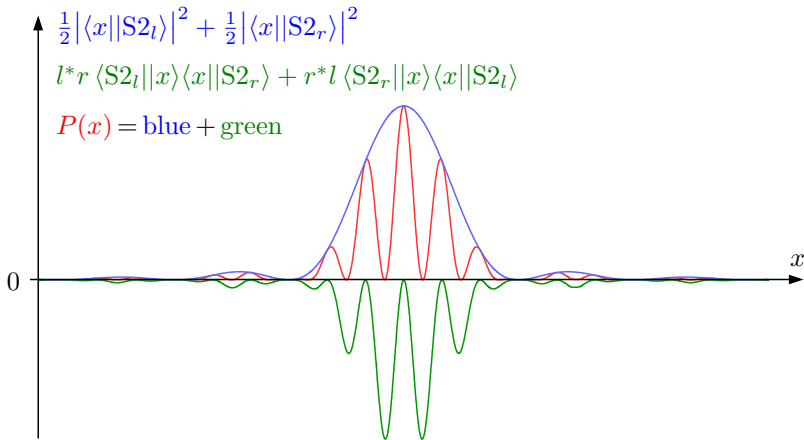


Fig. 5.6: The probability $P(x)$ to detect the neutron at position x

is three times their width.⁴⁶

It is worth taking a *very* close look at how the red curve is derived. The blue curve is the sum of the first and fourth terms in (5.18). With the modulus square $|\langle x | S_{2_l} \rangle|^2$ the neutron has uniquely come from the left slit to the detector position x ; with the modulus square $|\langle x | S_{2_r} \rangle|^2$ it has uniquely come from the right slit. These two terms are always positive and do not contribute to the interference between the two possibilities (left slit or right slit). The interference is described by the “mixed” terms — i. e. the second and third summands in (5.18) — in which both S_{2_l} and S_{2_r} appear. The value of these terms depends on the difference in path lengths from the left or right slit to the position x of the detector. If the difference in path lengths is equal to

$$0, \pm\lambda, \pm2\lambda, \pm3\lambda, \dots,$$

where λ is the wavelength of the neutron, then there is an interfer-

⁴⁶ For physicists: The blue line is $\sin^2(x)/x^2$, i. e. the square of the “slit function”. The red line is $\cos^2(3x) \sin^2(x)/x^2$.

ence-maximum at this point. But if the difference in path lengths is equal to

$$\pm\lambda/2, \pm3\lambda/2, \pm5\lambda/2, \dots,$$

then there is an interference minimum at this point; see fig. 2.4 on page 26. It is plausible that there is always a maximum at the point on the x -axis that is equidistant from both slits, i. e. at the center of the interference pattern; see also fig. 2.11 on page 40.

5.3 Measurements in Quantum Theory

This section is little more than a review of what has already been explained in the two preceding sections. As the topic of “measurement” is so important, however, it is worth to compile the key points in a clear summary.

The measuring device does not determine what state vector (or other properties) the quantum object under investigation had prior to the measurement; rather, it imposes one of its eigenvectors onto the quantum object. (5.19a)

An alternative manner of speaking is, that the measuring device prepares the object in one of its eigenstates. Every measuring device has only a limited set of eigenvectors that it can impose on the quantum object. In the example of the polarization measurement in the previous section (see fig. 5.5 on page 111) the photon was prepared in one of the two states $|x\rangle$ or $|y\rangle$, and the neutron was prepared by the detector in one of its 87 eigenstates $|x_1\rangle, |x_2\rangle, \dots, |x_{87}\rangle$.

Not every possible measurement result is equally likely. Rather, the probability of a specific measurement result is determined by Born’s rule (5.9):

Born's rule: A measuring device will impose onto a quantum object with probability

$$P_j = |\langle m_j || q \rangle|^2 \quad (5.19b)$$

its eigenvektor $|m_j\rangle$, i. e. prepare the object in state $|m_j\rangle$, if the object's state vector before the measurement is $|q\rangle$.

In the formalism of quantum theory, the eigenvectors of the measuring devices are elements of *vector spaces* that have exactly as many dimensions as the measuring device has eigenvectors. These are, of course, abstract mathematical spaces that have nothing to do with the three-dimensional space in which we are living. The two-dimensional space of the vectors $|x\rangle$ and $|y\rangle$ is easy to visualize and can also be drawn on paper, as shown e. g. in fig. 5.2 on page 104. But no one can really visualize an 87-dimensional space. The formal mathematical treatment of the abstract 87-dimensional space of the neutron detector's eigenvectors, however, is no more difficult than the formal mathematical treatment of the three-dimensional position space.

The eigenvectors of measuring devices are in quantum theory constructed such, that the following holds:

All eigenvectors of a measuring device are mutually orthogonal:

$$\begin{aligned} \langle m_j || m_k \rangle &= 0 & \text{if } k \neq j \\ \langle m_j || m_k \rangle &= 1 & \text{if } k = j \end{aligned} \quad (5.19c)$$

5.4 How Quantum Theorie was detected

Max Planck had discovered his "lucky guess interpolation formula" (3.3) for the blackbody radiation spectrum in 1900. For many years, this formula stood alone, with no apparent connection to the physical theories known at the time. The atomic model [30]

by Niels Bohr (1885–1962), in which electrons orbit around the atomic nucleus like planets around the sun, dated from 1913 and was likewise not a theory, but merely a model with many gaps, inconsistencies, and internal contradictions. Quantum theory was discovered only in 1925, but then twice!

The first to find it was Werner Heisenberg (1901–1976), then a 23-year-old assistant to Max Born (who later discovered Born's rule) in Göttingen. In early June 1925, Heisenberg was plagued by such a severe hay fever that he had to ask Born for time off and flee to the island Helgoland for ten days. There, he was able to work undisturbed on finding a better alternative to Bohr's atomic model.

While the orbits of electrons around the atomic nucleus played a central role in Bohr's model, they did not appear at all in the experimenters' observations. No one had ever seen an electron orbiting an atomic nucleus. Heisenberg hoped to arrive at better results if he removed all the unobserved baggage from the model and focused entirely on what the experimenters actually saw. What the experimenters actually saw were the positions of the spectral lines (at which wavelengths do atoms and molecules absorb light, and at which wavelengths do they not?), and their intensity (why are some spectral lines very strong, and others only very weak?).

Heisenberg succeeded in describing the simplest of all spectra — namely, the vibrational spectra of diatomic molecules — using arrangements of numbers based on a peculiar system. Heisenberg simply called them “number schemes”. When he returned to Göttingen with this work, Born was quite astonished. No one in atomic physics had ever come up with such an idea before. It turned out that Heisenberg's numerical schemes were something that mathematicians referred to as matrices. Heisenberg's novel quantum mechanics was therefore also called Matrix Mechanics.

In the months that followed, Heisenberg, Born, and Pascual

Jordan (1902–1980), another assistant of Born, further developed Matrix Mechanics. From the established theories of classical physics Matrix Mechanics differed primarily in four topics:

- * First, it was based on a non-commutative algebra. When two matrices A and B are multiplied, in general $AB \neq BA$.
- * Matrix Mechanics cannot be formulated using real numbers alone. Complex numbers are indispensable.
- * Since Newton’s time, differential equations have formed the backbone of all physical theories. Matrix Mechanics, on the other hand, is based on matrix transformations. This is an extremely laborious process compared to the elegant technique of differential equations.
- * Matrix Mechanics was a *completely* abstract mathematical formalism. No one could clearly visualize what was actually happening in quantum phenomena — not even Heisenberg.

Erwin Schrödinger (1887–1961), by then professor of theoretical physics at the University of Zürich, found Heisenberg’s matrix mechanics so “horrible”, that he devoted his Christmas break 1925/26 to the task of developing a “nicer” quantum theory. And he was successful! Unlike Heisenberg, Schrödinger had a clear picture of atomic processes in mind. He interpreted de Broglie’s wave theory of matter literally, and pondered what electron waves might look like within an atom.

When a musician plays the French horn, he can produce only a discrete set of “natural tones”, because the length of the horn must be an integer multiple of the wavelength of the sound wave. If he wants to play intermediate tones, he must stuff his fist into the bell of the horn to shorten the length of the tube accordingly. Just as the sound wave must fit into the tube of the French horn, so must the electron wave, which is bound to the atomic nucleus by electrostatic forces, fit into the space available to it.

Schrödinger succeeded to formulate this idea with mathematical precision, and to translate it into a differential equation. And as the solution of this differential equation — which came to be known as the Schrödinger equation — he indeed obtained the spectrum of the hydrogen atom. Correspondingly great was the general enthusiasm, which — however — soon gave way to abrupt disillusion.

When Schrödinger started to calculate the spectra of atoms with multiple electrons, he noticed that the electron waves do not oscillate in three-dimensional position space around the atomic nucleus, but rather in an abstract mathematical space that expands by three dimensions with each additional electron. As the abstract mathematical space for the hydrogen atom is only three-dimensional (because the hydrogen atom has only a single electron), Schrödinger had mistakenly confused it with three-dimensional position space.

Thus Quantum Theory remained quite abstract as well in Schrödinger's formulation. And the other distinctive features of Matrix Mechanics — namely, a non-commutative algebra and the necessity of complex numbers — also reappeared in Schrödinger's theory. In the spring of 1926, Schrödinger eventually realized that his “Wave Mechanics” and Heisenberg's “Matrix Mechanics” were ultimately mathematically completely equivalent. Regardless of whether one calculated a physical problem using Matrix Mechanics or Wave Mechanics, the result had to be the same due to the mathematical equivalence. It was as if Heisenberg and Schrödinger had written exactly the same text, only one in Chinese and the other in Japanese.

Regarding the content, both theories were equivalent. But formally, Schrödinger's differential equation formalism was far more elegant and efficient than Heisenberg's cumbersome matrix formalism. Therefore also the Göttingen physicists (everyone else

anyway) gratefully switched to Schrödinger's differential equation, while matrix mechanics vanished forever from the scene after barely one year.

In addition to those already mentioned (Heisenberg, Schrödinger, Born, Jordan), Wolfgang Pauli (1900–1958), then professor of theoretical physics at the University of Hamburg, and Paul Dirac (1902–1984), then a doctoral student at the University of Cambridge (UK), also played outstanding roles in the development of the nascent quantum theory. In Dirac's hands, Schrödinger's intuitive wave functions mutated into the abstract $|\text{state vectors}\rangle$ that we encountered in the first part of this chapter. And the elements of Heisenberg's matrices turned out to be $\langle\text{projection}||\text{amplitudes}\rangle$ onto these state vectors.

By mid 1926, physicists thus had a mathematical framework available, that produced a steady stream of physical results, all of which proved to be correct when tested experimentally. But no one felt to truly “understand” what was actually going on in quantum phenomena.

This was for the physicists a completely new experience. With all theories of classical physics, they had initially had a more or less clear idea of the underlying relationships, and then attempted to express these ideas in the most precise and simple mathematical form possible. With quantum theory, the situation was exactly opposite. The mathematical formalism lay ready-made on the table, but there was no intuitive explanation attached to it. Schrödinger's intuitive image of electron waves had turned out to be an error, and Heisenberg's number schemes had from the outset been a completely abstract mathematical construct.

Once again, it was Heisenberg who broke the deadlock in the stalled development. This time, however, he did not act alone, but together with Bohr, whose contribution to the “Copenhagen Interpretation of quantum theory” was at least as significant as

Heisenberg's. From 1924 to 1927, Heisenberg spent several months each year at Bohr's institute in Copenhagen. In the winter of 1926/27, the two physicists there together worked out the intuitive interpretation of the mathematical formalism of quantum theory, which later became known as the "Copenhagen Interpretation".

Heisenberg had begun studying physics at the University of Munich in the fall of 1920. Instead of first gaining a thorough understanding of classical physics, he immediately focused on theoretical atomic physics, which was taught there by Arnold Sommerfeld (1868–1951). Incidentally, due to this neglect of classical physics he nearly failed his final exam at the end of his studies and passed only with the lowest possible grade.

In early summer 1922, Sommerfeld traveled to Göttingen for a conference, at which Niels Bohr delivered several lectures on the current state of atomic physics. Heisenberg got the opportunity to accompany Sommerfeld. After a lecture in which Bohr discussed a paper he had published a few months earlier with his assistant Kramers, Heisenberg stood up and expressed doubts and objections. Heisenberg was confident in his position because he had recently presented on this article in Sommerfeld's seminar.

Bohr quickly recognized the extraordinary talent of the cheeky student and invited him to a long walk across the Hainberg on the outskirts of Göttingen, to discuss the matters at leisure. "This walk had the strongest influence on my later scientific development, or it would perhaps be better to say that my real scientific development started only on this walk", Heisenberg recalled decades later in his autobiography [31].

Heisenberg remembered Bohr to have said during this walk among other things: "In physics up to now, when one wanted to explain a new phenomenon, one could try, using existing concepts and methods, to trace the new phenomenon back to already known phenomena or laws. In atomic physics, however, we already know

that the existing concepts are certainly insufficient for this purpose. Consequently there can be no intuitive description of the structure of the atom, since such a description — precisely because it is supposed to be intuitive — would have to make use of the concepts of classical physics, which no longer capture the reality of what is going on. You understand that with such a theory, one is actually attempting something quite impossible. For we should say something about the structure of the atom, but we possess no language with which we could make ourselves understood.”

“If the internal structure of atoms is as inaccessible to an intuitive description as you say,” Heisenberg remembered to have asked in return, “and if we actually don’t have a language with which we could talk about these structures, will we then ever understand the atoms?” “Yes, we will,” so Bohr’s answer, “but in doing so, we will first have to learn what the word ‘understand’ does mean.”

The last sentence is characteristic of Bohr, and it is characteristic of the Copenhagen Interpretation of quantum theory. For, in essence, the analysis of the word “to understand” actually forms the core of this interpretation, on which I will comment at several places in the following chapters.

We will find this interpretation much easier to understand, and, above all, we will make significant progress on the topic of “particle location as a relational property”, if we get to know the “entanglement” of state vectors, which is the subject of the following chapter.

6 Entanglement

If a quantum system consists of several subsystems, then the state vector of the entire system may be “entangled”. In this chapter, we will clarify what exactly this means and what the consequences are. Schrödinger, who coined this term, described entanglement not as “*one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought.” [32] This is hardly exaggerated. We will not be able to eliminate the peculiarities of the phenomena we have dealt with in the preceding chapters through the analysis of entanglement; but we will be able to describe them more precisely and classify them more systematically than before.

This chapter is difficult because it is even more abstract than the previous ones. We won’t need any more mathematical methods of quantum theory than those introduced in the previous chapter — namely, state vectors and projection amplitudes. But we’ll be getting a heavy dose of both, the full brunt of it. Sorry, that can’t be avoided. Where intuition reaches its limits — and it certainly does when it comes to entanglement — only mathematics can help us further.

The effort is worthwhile because the analysis of entangled quantum systems yields three significant insights:

- * First, using entangled systems one can experimentally demonstrate that many properties, which we typically regard as properties of objects, are in fact relational properties. That is, they are not properties that a single object possesses “in and

of itself”, but rather properties generated by the interaction between that object and appropriate measuring devices.

- * Second, we will find evidence that nature acts non-locally in many quantum phenomena.
- * Third, the analysis of entangled quantum systems yields the proof — already announced in chapter 1 — that the results of measurements on quantum systems in individual cases⁴⁷ are strictly *irrational*, meaning that they are not determined by previous events and therefore cannot be calculated using (possibly yet undiscovered) laws of nature, but rather that “true” chance is at work here.

As usual, we will stick as closely as possible to key experiments that have been conducted on the topic of entanglement. To understand these experiments, we first need to clarify what the “spin” of electrons is.

6.1 Spin

In 1922, Otto Stern (1888–1969) and Walther Gerlach (1889–1979) conducted an experiment at the University of Frankfurt [33], which is sketched in fig. 6.1 on the next page:

In a vacuum chamber, a beam of silver atoms⁴⁸ — as indicated by the turquoise arrow — is directed at a slit which is 0.8 mm high and approximately 0.03 mm to 0.04 mm wide. 35 mm behind the slit, the atomic beam strikes a glass plate, onto which the silver is deposited.

In 6.1(a) the experiment is viewed from above, and 6.1(b) is showing the view from the front (in the direction of the atomic

⁴⁷ The statistical distribution of the results of a large number of measurements on identical quantum systems can be calculated very precisely, but not the result in an individual case. This is exactly like the results of a dice game.

⁴⁸ Ag is the chemical symbol for silver, from (Latin) *argentum* = silver.

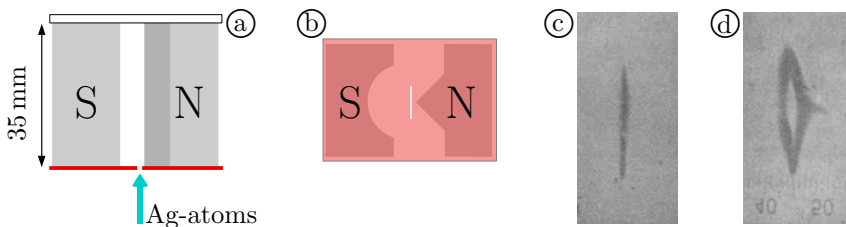


Fig. 6.1: The experiment of Stern and Gerlach

beam) onto the metal plate with the slit. This plate is drawn transparent in 6.1 (b) so that one can recognize the peculiar shape of the pole shoes of an electromagnet located between the slit and the glass plate. The north pole is shaped as a sharp edge, while a wide, round groove is milled into the south pole. As they travel from the slit to the glass plate, the silver atoms move closely along the edge of the north pole.

6.1 (c) shows the deposit of silver on the glass plate when the electromagnet is off, 6.1 (d) shows the deposit when the electromagnet is on. If you zoom into the photograph 6.1 (d) you can see at the bottom of the image the reversed and upside-down scale of the measuring eyepiece. 1 scale division = 0.05 mm, i. e. the distance from 40 to 50 is exactly 0.5 mm long.

The splitting of the silver-precipitate line in 6.1 (d) has nothing to do with diffraction and interference of atoms at the slit. This is evident simply from the fact that the splitting disappears when the electromagnet is switched off. The de Broglie wavelength of the silver atoms is also far too small to be detectable with such a wide slit.

Instead of presenting and commenting on the countless attempts at explanation proposed for the strange result of Stern and Gerlach's experiment in the following years, I jump straight to the year 1928. In that year, Paul Dirac [34, 35] introduced a "relativistically invariant" form of quantum theory, i. e. he had modified quantum

theory so that it was consistent with Einstein's special theory of relativity. (In the form discovered by Heisenberg and Schrödinger in 1925/26, Quantum theory did not yet do this.) In developing this theory, Dirac realized that electrons must be assigned a completely new property that does not exist in classical physics. This property is called *spin*. Associated with spin is a magnetic dipole moment of the electrons, which in an external magnetic field can always point only toward either the north pole or the south pole. Dirac's explanation of why the spin — and thus also the magnetic moment — of an electron can have only two directions is far (!!) too difficult for this book. Anyone who wants to know more about spin must study physics.⁴⁹

Because of the magnetic dipole moment⁵⁰ one can imagine silver atoms as tiny magnetic needles. But note: While the magnetic needle of a compass, which is turned by the Earth's magnetic field in the north-south-direction, can be turned with minimal effort into any other direction, the magnetic moment of the silver atoms is *always* aligned either *exactly* parallel or *exactly* antiparallel to an external magnetic field.

The Earth's magnetic field causes a compass needle to point in

⁴⁹ I warn unsuspecting readers against vivid images such as, for example, tiny rapidly spinning tops and the like, which appear time and again in popular science texts. Such images are, without exception, false and misleading. Spin is a highly abstract concept that can be precisely described in mathematical terms but cannot be linked to images that fit into the human brain. Just one example: If you rotate a compass needle by 360° , it is identical to a compass needle that has not been rotated at all. In contrast, the spin (and thus also the magnetic moment) of an electron is only identical to that of an electron spin that has not been rotated at all after a rotation of 720° (but not after a rotation of 360°).

⁵⁰ For physicists: This magnetic moment is caused by the electron shell of the silver atom in its ground state. Its atomic nucleus has an additional magnetic moment, but this is more than ten thousand times smaller than the magnetic moment of the electron shell, and can therefore be safely ignored when discussing the Stern-Gerlach experiment.

the north-south direction, but it does not cause the entire needle to move toward the north pole or the south pole. This is because the north pole of the needle is attracted to the Earth's magnetic south pole with the same force as the south pole of the needle is attracted to the Earth's magnetic north pole.⁵¹ This changes, however, if the magnetic field is not equally strong everywhere, but stronger at one end of the compass needle than at the other.

Precisely this is what Stern and Gerlach achieved through the unique shape of their magnet's poles, as illustrated in fig. 6.1 on page 128. Near the sharply crafted north pole, the magnetic field is much stronger than on the other side of the beam, toward the south pole. When the north pole of the magnetic dipole moment of a silver atom is directed toward the north pole of the magnet, the north pole of the atom is repelled more strongly by the north pole of the magnet than the south pole of the atom is repelled by the south pole of the magnet. Thus, the atom as a whole is deflected toward the south pole of the magnet. If, on the other hand, the south pole of the magnetic dipole moment of a silver atom is directed toward the north pole of the magnet, then the atom is deflected toward the north pole of the magnet.

If the magnetic moments of the silver atoms were oriented in arbitrary directions relative to the electromagnet's field, then the narrow line in fig. 6.1(c) would broaden into a wide spot when the magnetic field is applied. In fact, however, fig. 6.1(d) shows a splitting into two clearly separated lines. Close to the sharp edge, the deflection of one line is particularly pronounced; at a greater distance from the edge, the magnetic field is not inhomogeneous enough to separate the two lines. If, instead of a broad spot, there

⁵¹ Like magnetic poles repel each other; unlike magnetic poles attract each other. The Earth's magnetic north pole is currently located near the geographic north pole in northernmost Canada and is moving toward Siberia at a speed of about 50 km/year. The south pole of a compass needle points toward the north pole of the Earth's magnetic field.

are only two clearly separated lines, this means that the magnetic moments of the silver atoms are not oriented in arbitrary directions in space, but rather have one or the other of only two possible directions relative to the field of the electro-magnet.

In fig. 6.2, the turquoise arrow represents silver atoms that have passed through a Stern-Gerlach-magnet oriented in α -direction. We assign the state vector $|\uparrow\rangle_\alpha$ to the atoms that have been deflected toward the north pole, and we assign the state vector $|\downarrow\rangle_\alpha$ to the atoms that have been deflected toward the south pole. The silver atom has 47 electrons, but 46 of them form 23 pairs $\uparrow\downarrow$ that magnetically **compensate** each other. Therefore only a single arrow appears in the state vector of silver atoms.

The second Stern-Gerlach-magnet in fig. 6.2 is a measuring device with the two eigenvectors $|\uparrow\rangle_\beta$ and $|\downarrow\rangle_\beta$. The probability that a silver atom, which was deflected by the first magnet toward N, will also be deflected by the second magnet toward N is, according to Born's rule,

$$P(\uparrow_\alpha, \uparrow_\beta) \stackrel{(5.19b)}{=} |\beta\langle\uparrow|\uparrow\rangle_\alpha|^2.$$

By conducting large series of measurements both with atoms deflected to the N direction by the first magnet and with atoms

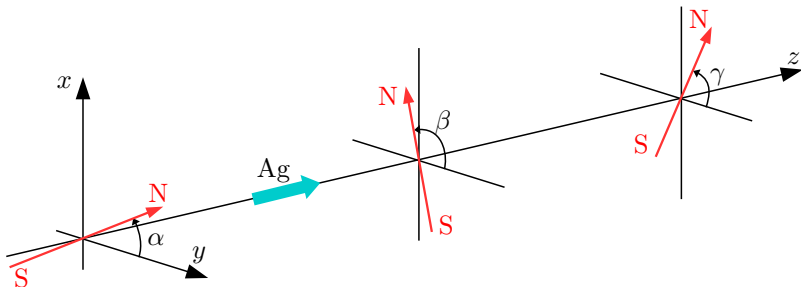


Fig. 6.2: Measurement with several Stern-Gerlach-magnets

deflected to the S direction by the first magnet, and by varying the angles of the magnets, one finds⁵² the following values for the various probabilities:¹⁴

$$P(\uparrow_\alpha, \uparrow_\beta) = |\beta \langle \uparrow \uparrow \uparrow \rangle_\alpha|^2 = \cos^2(\beta/2 - \alpha/2) \quad (6.1a)$$

$$P(\uparrow_\alpha, \downarrow_\beta) = |\beta \langle \downarrow \uparrow \uparrow \rangle_\alpha|^2 = \sin^2(\beta/2 - \alpha/2) \quad (6.1b)$$

$$P(\downarrow_\alpha, \uparrow_\beta) = |\beta \langle \uparrow \uparrow \downarrow \rangle_\alpha|^2 = \sin^2(\beta/2 - \alpha/2) \quad (6.1c)$$

$$P(\downarrow_\alpha, \downarrow_\beta) = |\beta \langle \downarrow \uparrow \downarrow \rangle_\alpha|^2 = \cos^2(\beta/2 - \alpha/2) \quad (6.1d)$$

If the same experiment is performed to measure the deflection of the atoms by the third magnet, one finds:

$$P(\uparrow_\beta, \uparrow_\gamma) = |\gamma \langle \uparrow \uparrow \uparrow \rangle_\beta|^2 = \cos^2(\gamma/2 - \beta/2) \quad (6.1e)$$

$$P(\uparrow_\beta, \downarrow_\gamma) = |\gamma \langle \downarrow \uparrow \uparrow \rangle_\beta|^2 = \sin^2(\gamma/2 - \beta/2) \quad (6.1f)$$

$$P(\downarrow_\beta, \uparrow_\gamma) = |\gamma \langle \uparrow \uparrow \downarrow \rangle_\beta|^2 = \sin^2(\gamma/2 - \beta/2) \quad (6.1g)$$

$$P(\downarrow_\beta, \downarrow_\gamma) = |\gamma \langle \downarrow \uparrow \downarrow \rangle_\beta|^2 = \cos^2(\gamma/2 - \beta/2) \quad (6.1h)$$

Obviously the result of the previous measurement is completely overwritten by the result of the new measurement. In the last four equations, the angle α of the first magnet does no more appear at all.

6.2 Measurement and Reality

The term “measurement” has a different meaning in quantum theory than in classical physics. In classical physics, a measurement determines a property of an object, that the object already possessed before the measurement, and continues to possess after the measurement. In quantum theory, the situation is different: the

⁵² These probabilities can not only be derived from experiments but also calculated theoretically. In this book I try, however, to spare the readers wherever possible mathematical calculations.

object's state vector remains unchanged during the measurement only if it was already prior to the measurement identical to one of the eigenvectors of the measuring device. Otherwise, the measuring device imposes one of its eigenvectors onto the object. One says: The object is prepared in one of the measuring device's eigenstates. Which one? That is determined by chance. According to Born's rule (5.19b), one can calculate the probability that the object is prepared in a specific eigenstate of the measuring device by projecting the state vector of the object (before the measurement) onto this eigenstate of the measuring device. The square of the projection amplitude is the probability sought.

In general, the object being measured is thus altered by the measurement. In the measurement of the direction (\uparrow or \downarrow) of the silver atom's magnetic moment relative to the direction of the external magnetic field, using the three measuring devices (i. e. the three Stern-Gerlach-magnets) as illustrated in fig. 6.2, the atom is first prepared in the state $|\uparrow\rangle_\alpha$ or $|\downarrow\rangle_\alpha$, then in the state $|\uparrow\rangle_\beta$ or $|\downarrow\rangle_\beta$, and finally in the state $|\uparrow\rangle_\gamma$ or $|\downarrow\rangle_\gamma$. The measurement appears to be a creative act: with each measurement, a new state is created, and at the same time, the previous state is annihilated.

Heisenberg [36] had already in spring 1927 pointed out a very similar situation regarding the measurement of an object's position and momentum. By means of various examples, he demonstrated that a precise measurement of an object's position inevitably causes as a side effect a change in its momentum, the uncertainty of which increases the more precise the position measurement is. And that a measurement of momentum inevitably causes as a side effect a change in position, the uncertainty of which increases the more precise the momentum measurement is. Consequently, one cannot measure the position and momentum of an object simultaneously with arbitrary precision, but must make compromises. Heisenberg estimated that the product of the inevitable uncertainties in posi-

tion and momentum measurements is at least roughly as large as Planck's constant:

$$\left(\begin{array}{c} \text{inaccuracy of} \\ \text{position measurement} \end{array} \right) \cdot \left(\begin{array}{c} \text{inaccuracy of} \\ \text{momentum measurement} \end{array} \right) \geq h \quad (6.2)$$

In section 8.1 I will discuss this topic in more detail.

The question that was hotly debated at the time was: Do position and momentum actually have precise values at the same time, which only cannot be determined simultaneously because of the restricted measurement capabilities? And does the magnetic moment of silver atoms have precise values (\uparrow or \downarrow) in different directions at the same time, which we simply cannot measure simultaneously because the field of a Stern-Gerlach-magnet can have only one direction at a time? Or must the measurement of an object's momentum and position be regarded as a creative process, such that, due to (6.2), a precise position and a precise momentum of an object can never be created simultaneously, and therefore, in reality, the position and momentum of an object can never have precisely determined values at the same time? Does the magnetic moment of a silver atom in reality always exist only in the direction of the magnet with which it was last measured, but not in other directions?

Quantum theory assumes the latter. According to this theory, the magnetic moment of a silver atom exists only in the direction of the magnetic field of the magnet with which it was last measured, and the positions and momenta of arbitrary objects never have simultaneously more precise values than consistent with (6.2).

6.3 EPR = Einstein, Podolski, Rosen

Albert Einstein disagreed. In his view, Heisenberg's relation (6.2) did indeed correctly describe the limited possibility to measure position and momentum simultaneously. But he was convinced

that in reality — even though it is only partially accessible for measuring instruments — these properties do have exact values at all times. In 1935, he had an idea of how this could be proven — at least in principle. Together with his assistants Boris Podolski (1896–1966) and Nathan Rosen (1909–1995), he published this idea under the title “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” [37]. It was the last of Einstein’s numerous and important contributions to quantum theory, and the only one that contained an error — or, to put it more clearly: the only one in which Einstein was completely off the mark. As this article is cited and discussed incredibly frequently in the physics literature, it has become customary to abbreviate the names of the three authors as EPR.

The question of the completeness of quantum theory was merely a [rhetorical](#) one. In their article, EPR left no doubt that they considered quantum theory to be incomplete, specifically incomplete with regard to properties of quantum objects to which quantum theory, strangely enough, assigns only imprecise values or no values at all, because these values can only imprecisely or not at all be measured.

In their paper, EPR sought to prove that properties such as the positions and momenta of particles, or the magnetic moments \uparrow or \downarrow of silver atoms in any given direction, do indeed have precisely defined values at all times, even if these values have not (yet) been measured or cannot be measured (directly). Curiously, however, it was just the analysis of their argument that ultimately led to the experimental proof that this is indeed the case: the magnetic moments \uparrow or \downarrow of silver atoms in certain directions, or the positions and momenta of quantum objects, are only created by the measurement, and are therefore truly and actually defined only with the limited precision that the measurements allow. Before the measurement, they are not only unknown, but they simply do

not exist.

EPR examined this issue based on the incompatibility of precise measurements of position and momentum, as noted in (6.2). But I will explain their argument by an experiment described by David Bohm (1917–1992) in his textbook on quantum theory [38]. In the form discussed by Bohm, this experiment has never been carried out. The great advantage of Bohm’s “thought experiment” is, however, that it is much simpler in design and therefore much easier to understand than the experiments actually conducted in the laboratory many decades later, which we will discuss in chapter 7.

Bohm considered two not further specified atoms D and E, which are deflected like silver atoms by Stern-Gerlach-magnets toward N or S. He also assumed that these two atoms combine to form a DE-molecule in such a way that this molecule is not deflected at all when passing through a Stern-Gerlach-magnet, neither toward N nor toward S. I mentioned earlier that silver atoms have 47 electrons, 46 of which arrange themselves in pairs $\uparrow\downarrow$ so that their magnetic moments **compensate**. Consequently it is entirely plausible to assume that the atoms D and E in the molecule DE arrange themselves in such a way that their magnetic moments cancel each other out, and the molecule is magnetically

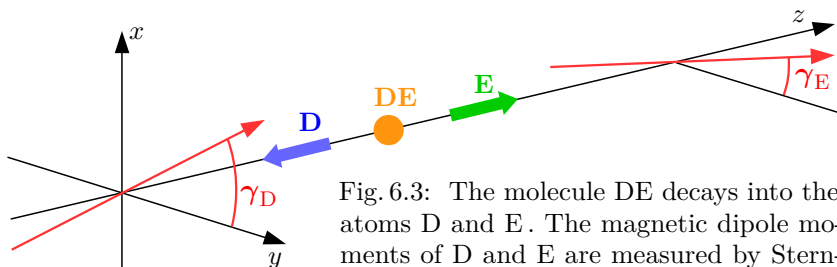


Fig. 6.3: The molecule DE decays into the atoms D and E. The magnetic dipole moments of D and E are measured by Stern-Gerlach-magnets that are rotated by angles γ_D resp. γ_E relative to the y -axis.

neutral.

Bohm's third assumption was that the DE molecule is unstable, and decays after some time back into the atoms D and E, which are then analyzed using Stern-Gerlach-magnets as sketched in fig. 6.3. I should add now that spin (which gives rise to the magnetic moments of the atoms) is a conserved quantity. Spin cannot simply appear out of nowhere, nor can it simply vanish. In the DE molecule, the spins of the two atoms were arranged antiparallel $\uparrow\downarrow$, so that the magnetic moments canceled each other out to zero, and the total spin of the molecule was zero. Then, due to conservation of spin, the total spin of the flying-off atoms must still be zero, i. e. the magnetic moments of the two atoms must remain aligned antiparallel $\uparrow\downarrow$. Therefore, after the molecule's decay, either atom D is in the \downarrow state and atom E in the \uparrow state, or atom D is in the \uparrow state and atom E in the \downarrow state.

Which of these two possibilities actually holds true? Strangely enough, quantum theory leaves this question open. In quantum theory, the overall system D & E of the two atoms D and E approaching the Stern-Gerlach-magnets (see fig. 6.3) is described prior to measurement by the state vector

$$|D\&E\rangle = r |\downarrow\rangle_D |\uparrow\rangle_E + s |\uparrow\rangle_D |\downarrow\rangle_E \quad (6.3)$$

$$\text{with } |r|^2 = |s|^2 = \frac{1}{2} .$$

This type of state vectors is referred to as “entangled”. It is characteristic for entangled state vectors, that they are not the product of state vectors of subsystems. Let's try to construct (6.3) as the product of the state vectors

$$|D\rangle = q |\uparrow\rangle_D + u |\downarrow\rangle_D \quad \text{with } |q|^2 + |u|^2 = 1 \quad (6.4a)$$

$$|E\rangle = v |\uparrow\rangle_E + w |\downarrow\rangle_E \quad \text{with } |v|^2 + |w|^2 = 1 \quad (6.4b)$$

of the atoms D and E:

$$\begin{aligned}
|D\rangle|E\rangle &= \left(q|\uparrow\rangle_D + u|\downarrow\rangle_D \right) \left(v|\uparrow\rangle_E + w|\downarrow\rangle_E \right) = \\
&= qv|\uparrow\rangle_D|\uparrow\rangle_E + qw|\uparrow\rangle_D|\downarrow\rangle_E + \\
&\quad + uv|\downarrow\rangle_D|\uparrow\rangle_E + uw|\downarrow\rangle_D|\downarrow\rangle_E \tag{6.5} \\
&\text{with } |q|^2 + |u|^2 = 1 \text{ and } |v|^2 + |w|^2 = 1
\end{aligned}$$

To reproduce from (6.5) the entangled state (6.3), we would need

$$qw = s \implies q \neq 0 \text{ and } w \neq 0 \tag{6.6a}$$

$$uv = r \implies u \neq 0 \text{ and } v \neq 0 \tag{6.6b}$$

and at the same time

$$qv = 0 \implies q = 0 \text{ or } v = 0 \tag{6.6c}$$

$$uw = 0 \implies u = 0 \text{ or } w = 0 . \tag{6.6d}$$

It is impossible to meet the four conditions (6.6) at the same time. Consequently the entangled state vector (6.3) can not be constructed as the product of the state vectors (6.4) of the single atoms.

What, then, are the state vectors of the individual atoms if the state vector of the entire system is given by (6.3)? Well, (6.4) are indeed the most general state vectors that can be defined for the single atoms. If (6.3) cannot be written as a product of these states, then this means not more nor less than that, in the case of (6.3), quantum theory does not assign any state vectors at all to the single atoms.

While the state vectors $|\uparrow\rangle_D$, $|\downarrow\rangle_D$, $|\uparrow\rangle_E$, $|\downarrow\rangle_E$ of the two single atoms are contained in the entangled state vector (6.3), the single atoms have, so to speak, no independent existence. In the formalism of quantum theory, they exist only as components of the overall system D & E.

Since atoms D and E are contained within the overall system D & E but do not exist as independent entities, it follows that

certain properties — such as magnetic moments \uparrow or \downarrow along any axes — do not exist either. This is not merely a bold assertion of quantum theory; rather, it has been experimentally proven that this is indeed the case. All the lengthly discussions in this chapter serve only to enable the readers to understand the experimental evidence.

Quantum theory does not assign the two atoms distinct magnetic moments \uparrow or \downarrow along specific spatial axes, but it specifies a correlation: The magnetic moments present in the overall system D&E must cancel each other out to zero. These products appear in the entangled state vector (6.3):

$$|\uparrow\rangle_D |\downarrow\rangle_E \quad \text{and} \quad |\downarrow\rangle_D |\uparrow\rangle_E \quad (6.7a)$$

These product can mutually compensate to zero only, if

$$|\uparrow\rangle_D = |\uparrow\rangle_E \quad \text{and} \quad |\downarrow\rangle_D = |\downarrow\rangle_E . \quad (6.7b)$$

Since we describe the overall system of the two atoms by a common state vector $|\text{D\&E}\rangle = (6.3)$, we must also regard the two Stern-Gerlach magnets used to analyze the atoms (see fig. 6.3) as a single measuring device with four eigenvectors:

$$|\uparrow\rangle_{\gamma_D} |\uparrow\rangle_{\gamma_E} , |\uparrow\rangle_{\gamma_D} |\downarrow\rangle_{\gamma_E} , |\downarrow\rangle_{\gamma_D} |\uparrow\rangle_{\gamma_E} , |\downarrow\rangle_{\gamma_D} |\downarrow\rangle_{\gamma_E} \quad (6.8)$$

The probability P of each of the four possible measurement outcomes is according to Born's rule (5.19b) equal to the square of the modulus of the projection amplitude of the measuring device's corresponding eigenvector onto the state vector $|\text{D\&E}\rangle = (6.3)$ of the measured object. For example, let us calculate the probability of the result $|\uparrow\rangle_{\gamma_D} |\uparrow\rangle_{\gamma_E}$:

$$\begin{aligned}
P(\uparrow_{\gamma_D} \uparrow_{\gamma_E}) &= \left| \left(\gamma_D \langle \uparrow | \gamma_E \langle \uparrow | \right) |6.3\rangle \right|^2 = \\
&= \left| \left(\gamma_D \langle \uparrow | \gamma_E \langle \uparrow | \right) \left(r | \downarrow \rangle_D | \uparrow \rangle_E + s | \uparrow \rangle_D | \downarrow \rangle_E \right) \right|^2 = \\
&= \left| r \gamma_D \langle \uparrow || \downarrow \rangle_D \gamma_E \langle \uparrow || \uparrow \rangle_E + s \gamma_D \langle \uparrow || \uparrow \rangle_D \gamma_E \langle \uparrow || \downarrow \rangle_E \right|^2 \quad (6.9)
\end{aligned}$$

These projection amplitudes differ from the projection amplitudes whose values have been indicated in (6.1):

$$\begin{aligned}
|\gamma \langle \uparrow || \uparrow \rangle_\beta|^2 &= |\gamma \langle \downarrow || \downarrow \rangle_\beta|^2 \stackrel{(6.1)}{=} \cos^2(\gamma/2 - \beta/2) \\
|\gamma \langle \downarrow || \uparrow \rangle_\beta|^2 &= |\gamma \langle \uparrow || \downarrow \rangle_\beta|^2 \stackrel{(6.1)}{=} \sin^2(\gamma/2 - \beta/2)
\end{aligned}$$

These projection amplitudes always specify the orientation (β or γ) of the magnetic field that, together with the silver atom, has created the magnetic moment \uparrow or \downarrow along this axis. This information is missing in the state vectors of atoms D and E in (6.9). This is meant seriously. As long as the direction of the magnetic moments of D and E has not yet been created by a magnetic field, they have no direction according to quantum theory. (6.9) can still be computed by writing the state vectors of the atoms in the form

$$| \uparrow \rangle_D \stackrel{(6.7)}{=} | \uparrow \rangle_E = | \uparrow \rangle_{\alpha+\gamma_D} = e^{i\alpha} | \uparrow \rangle_{\gamma_D} \quad (6.11a)$$

$$| \downarrow \rangle_D \stackrel{(6.7)}{=} | \downarrow \rangle_E = | \downarrow \rangle_{\alpha+\gamma_D} = e^{i\alpha} | \downarrow \rangle_{\gamma_D} \quad (6.11b)$$

with arbitrary angle α .

The angle γ_D of the Stern-Gerlach magnet is shown in fig. 6.3. The complex function $e^{i\alpha}$ rotates the eigenvector $| \uparrow \rangle_{\gamma_D}$ of this magnet by an additional angle α , i. e. $| \uparrow \rangle_{\alpha+\gamma_D}$ is the eigenvector of a Stern-Gerlach magnet rotated by the angle $\gamma + \alpha$. The key point of this notation is that the direction of this (purely imaginary, in reality not actually existing) magnet is, due to “ α arbitrary”, just as undefined as the directions of the vectors $| \uparrow \rangle_D$ and $| \uparrow \rangle_E$. I know

that only physicists and mathematicians can truly understand (6.11). All other readers should simply trust that this notation is correct and reasonable.

Inserting (6.11) into (6.9), we get

$$\begin{aligned}
 P(\uparrow_{\gamma_D} \uparrow_{\gamma_E}) &= \left| r e^{i\alpha^2} \overbrace{\langle \uparrow \parallel \downarrow \rangle_{\gamma_D}}^{=0} \gamma_E \langle \uparrow \parallel \uparrow \rangle_{\gamma_D} + \right. \\
 &\quad \left. + s e^{i\alpha^2} \overbrace{\langle \uparrow \parallel \uparrow \rangle_{\gamma_D}}^{=1} \gamma_E \langle \uparrow \parallel \downarrow \rangle_{\gamma_D} \right|^2 = \\
 &= \overbrace{|s|^2}^{=1/2} \cdot \underbrace{|e^{i\alpha^2}|^2}_{=1} \cdot \left| \gamma_E \langle \uparrow \parallel \downarrow \rangle_{\gamma_D} \right|^2 \stackrel{(6.1)}{=} \frac{1}{2} \sin^2(\gamma_E/2 - \gamma_D/2) \quad (6.12a)
 \end{aligned}$$

In the same way, we find

$$P(\uparrow_{\gamma_D} \downarrow_{\gamma_E}) = \frac{1}{2} \cos^2(\gamma_E/2 - \gamma_D/2) \quad (6.12b)$$

$$P(\downarrow_{\gamma_D} \uparrow_{\gamma_E}) = \frac{1}{2} \cos^2(\gamma_E/2 - \gamma_D/2) \quad (6.12c)$$

$$P(\downarrow_{\gamma_D} \downarrow_{\gamma_E}) = \frac{1}{2} \sin^2(\gamma_E/2 - \gamma_D/2) . \quad (6.12d)$$

These probabilities are plotted in fig. 6.4 on the next page as a function of $(\gamma_E - \gamma_D)$. Note that $P(\downarrow_D \downarrow_E) = P(\uparrow_D \uparrow_E)$ and $P(\downarrow_D \uparrow_E) = P(\uparrow_D \downarrow_E)$, and that the probabilities depend only on the difference $(\gamma_E - \gamma_D)$ between the two angles, but not on their absolute values. The sum of the four probabilities is equal to 1 for every angle, as it must be.

Now we place the Stern-Gerlach magnet, which measures the magnetic moment of atom E, at a *much* greater distance from the point where the molecule DE decays than the magnet that measures the magnetic moment of D. Thus, we will already know the result of the measurement of D while the atom E is still on the

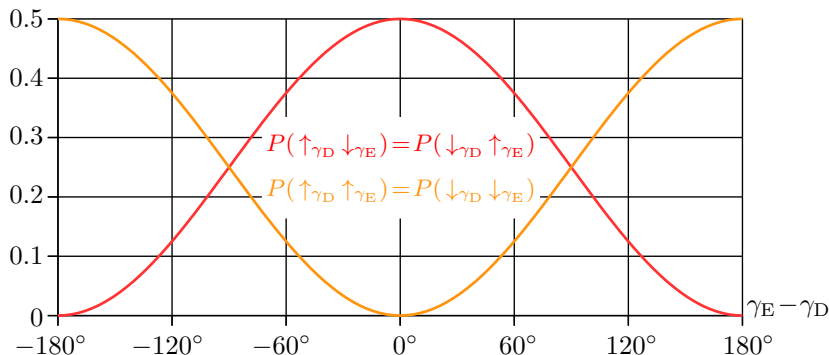


Fig. 6.4: The four probabilities (6.12)

way to “its” magnet. If the two magnets are oriented identically ($\gamma_E = \gamma_D$), then — based on (6.12) and fig. 6.4, we know with certainty that E will be measured as \downarrow (i. e. will be deflected toward S) if D was measured as \uparrow (i. e. was deflected toward N), and that E will be measured with \uparrow (i. e. will be deflected toward N) if D was measured with \downarrow (i. e. was deflected toward S). We know this with certainty, even if E has not yet reached its magnet.

Suppose atoms D and E reach their respective Stern-Gerlach-magnets at such times that not even a radio signal traveling at the speed of light could transmit the result of the measurement on atom D to atom E before the measurement on atom E is completed, and vice versa.

To increase the difficulty, corresponding experiments were later conducted in such a way that the settings of the measuring instruments (in our example, the angles γ_D and γ_E) were altered by random number generators so shortly before the arrival of the atoms that atom D could not know which detector-setting atom E would find, and vice versa, unless some mysterious communication at superluminal speed were taking place. Even under such difficult conditions, the correlations (6.12) were realized without error in

all experiments.

Without even coming close to atom E with a measuring device, we can thus predict the result of the measurement of E for any angle γ_E with certainty, simply by setting $\gamma_D = \gamma_E$ and noting the result of D. How do the atoms do this if their magnetic moment has no direction before the measurement? For example, if atom D finds its magnet set to $\gamma_D = 47^\circ$ and chooses \uparrow at the moment of measurement, then atom E must choose \downarrow if the random generator has also set its magnet to $\gamma_E = 47^\circ$.

EPR [37] were convinced that there were only two possible explanations for the correlation (6.12):

Either — as assumed by quantum theory — the magnetic moments \uparrow or \downarrow of atoms D and E do not yet exist in the direction of any spatial axes as long as they have not been measured, but are created only in the moment of measurement. Under this assumption, the correlation (6.12) can only be realized if the setting γ_D of the Stern-Gerlach-magnet used to measure D and the result obtained during the measurement is instantly transmitted — faster than with the speed of light, without the slightest delay — in some unknown manner to the location where the direction of the magnetic moment of E is measured, and influences this measurement such that the two measurement results are correlated according to (6.12). Communication at superluminal speed, however, was considered by EPR an out-of-the-question idea.

Alternatively, the magnetic moments of atoms D and E have — relative to any spatial axes — specific orientations \uparrow or \downarrow from the very moment the DE molecule decays. Because the antiparallel $\uparrow\downarrow$ orientation of the magnetic moments was preserved during the decay of DE, the magnetic moments of D and E are, of course, correlated along any spatial axes according to (6.12). During the measurement, the magnetic moments along these spatial axes are not generated anew; rather, the magnetic moments that have

existed since the decay of DE are simply detected.

For future reference, we restate the two alternative explanations in more general terms:

A Many properties of quantum objects do not yet exist until they have been measured; rather, they are created only at the moment of measurement. Under this assumption, the correlations between the properties of entangled subsystems can only be realized if the settings of the measuring apparatus and the results obtained during the measurement are instantly — faster than with the speed of light, without the slightest delay — transmitted in some unknown way between the locations where the measurements take place, thereby bringing about the correlation of the measurement results.

B The subsystems possess the properties, that will later be measured, already before the measurement takes place. During the measurement, these properties are not created; rather, the measurement merely detects the properties that already existed before.

EPR were convinced that only alternative **B** could be correct. With this alternative, the correlations (6.12) find a clear, simple, and reasonable explanation. Quantum theory, however, predicts prior to measurement merely the correlation (6.12); it does not predict whether atom D will be measured as \uparrow and atom E as \downarrow , or atom D as \downarrow and atom E as \uparrow , and it certainly does not predict these results for arbitrary orientations of the Stern-Gerlach magnets. For this reason, EPR deemed quantum theory incomplete.

The EPR argument was met with approval by some physicists, but with rejection by far more physicists. Bohr firmly refused this point of view. A few weeks after the publication of the article by Einstein, Podolski, and Rosen, his answer [39] was published in the

same journal. This article is a prime example for Bohr's notorious, convoluted, and extremely difficult-to-understand writing style. One must read the article at least three times with great patience to figure out, partly by reading between the lines, what he actually wants to say. The most important sentence can be found on the second of the seven pages:

“The impossibility of a closer analysis of the reactions between the particle and the measuring instrument is indeed no peculiarity of the experimental procedure described, but is rather an essential property of any arrangement suited to the study of the phenomena of the type concerned, where we have to do with a feature of *individuality* completely foreign to classical physics.” [39]

Bohr had the word *individuality* printed in italics to emphasize it, and he left no doubt — neither here nor in his other writings — that he meant the term literally. (Latin) *dividere* = to divide, (Latin) *in-* = un- (negation). Individuality thus means something like undividedness or wholeness.

According to Bohr, quantum phenomena can only be adequately understood through a holistic description. They can only be broken down into parts due to interaction with the environment, e. g. due to interaction with a measuring instrument. Until that is done, it is not meaningful to speak of the parts and their properties.

In his view, alternative B was completely wrong, but A also did not really capture the essence of the matter, because it refers to communication between parts of the system D & E, and thus fails to properly account for the holistic nature of quantum phenomena.

According to Bohr's understanding, the notion that atoms D and E are two independent objects during their flight toward their respective Stern-Gerlach magnets is incorrect. Rather, the entire system D & E constitutes a *single* quantum phenomenon that must not even mentally be divided into parts, no matter how far apart atoms D and E have moved. In the individual quantum

phenomenon D & E, the magnetic moments $\uparrow\downarrow$ of the two atoms cancel each other out to zero. But the atoms D and E do not exist as independent objects. Therefore, neither the magnetic moment of D nor the magnetic moment of E exists, and certainly these moments do not exist in the direction of any spatial axes. The lifetime of the quantum phenomenon D & E ends only when the measurement with the Stern-Gerlach-magnets creates two new quantum phenomena, namely the independent atoms D and E with magnetic moments $|\uparrow\rangle_{\gamma_D}$ or $|\downarrow\rangle_{\gamma_D}$ and $|\uparrow\rangle_{\gamma_E}$ or $|\downarrow\rangle_{\gamma_E}$ along the field directions of the two magnets.

Let's formulate Bohr's point of view as alternative C in general form:

- C** The idea that the parts of an entangled system are independent objects is incorrect. Rather, the system as a whole constitutes an *individual* quantum phenomenon that must not even mentally be broken down into parts, no matter how far apart its components are. The individual quantum phenomenon has certain properties. Its components do not exist as independent objects, however, and therefore do not possess all the properties that the overall system has. The lifetime of the individual quantum phenomenon ends only when a measurement creates two new quantum phenomena from it — namely, the independent subsystems along with their properties.

Bohr's perspective was not simply plucked out of thin air. Rather, quantum physicists had recognized in their analysis of quantum theory — which had been discovered in 1925/26 through ingenious guesswork — that this theory describes phenomena holistically, at least far more holistically than one was accustomed to in classical physics. It is precisely this holistic aspect that is visible in the entangled state vector (6.3): In quantum theory, a state vector is assigned only to the entire system D & E, but not to the individual

components D and E, and this state vector determines a holistic property, namely the correlation (6.12).

Bohr's considerations on

6.4 (In-)Divisible Phenomena

are so important, that I dedicate here an extra section to them. Fig. 6.5 shows a wire pendulum: Suspended from a thin wire of length L , which in turn is attached to a fixed point \ast , the black sphere swings back and forth just above the Earth's surface. If the weight of the wire is negligible compared to the weight of the black sphere, and if the mass density of the sphere is so high that air resistance can be neglected, then

the frequency $\nu = \sqrt{g/L}/(2\pi)$ of the pendulum can be calculated using the known gravitational acceleration $g = 9.81 \text{ m/s}^2$ near the Earth's surface.

This is already a quite useful physical description of the sphere's motion, but it can definitely be made still more precise. The most obvious improvement would be to take into account how the sphere's motion is slowed down by the air. One could increase the accuracy even further by taking into account not only the gravitational attraction of the Earth but also that of the Moon, or the gravitational attraction exerted by the people standing around the pendulum and watching it swing, or even the gravitational attraction of all the planets and the Sun, or — to make it truly perfect — the gravitational attraction of every object in the universe. If one wanted to achieve a “truly perfect” description of the pendulum, then “in principle” one would also have to take these tiny contributions into account. For the range of gravity (and also

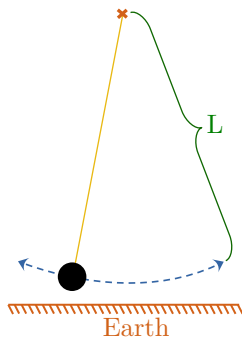


Fig. 6.5: A wire-pendulum

the range of the electromagnetic interaction) is — according to everything physics knows by today — actually infinite.

Of course, nobody does that, because nobody can measure with such precision as to detect the minuscule gravitational influence of distant galaxies. To avoid unnecessary effort, we ignore not only all immeasurably small factors but also — depending on the desired level of accuracy — quite noticeable side effects such as the braking effect of air. On the other hand, there also are details that must not be neglected under any circumstances if one wishes to achieve an at least roughly correct description of the sphere's motion; these are the wire, the fixed point to which the wire is attached, and the Earth with its gravitational effect. Thus, what is depicted in fig. 6.5 constitutes the indivisible phenomenon “wire pendulum”. If one were to consider the sphere, the wire, the fixed point, and the Earth separately, it would be impossible to arrive at a result that bears even a remote resemblance to the oscillating motion of a pendulum.

The molecule DE, sketched in fig. 6.3 on page 136, which decays into the phenomenon D & E consisting of the flying-away atoms D and E, is also an “indivisible phenomenon”, whose entangled state vector has been given in (6.3). If one were to consider the atoms D and E separately, the correlation between them — namely, the antiparallel $\uparrow\downarrow$ arrangement of their spins (resp. their magnetic moments) — would no longer be discernible. There is a fundamental difference, however, between the indivisible phenomenon “wire pendulum” and the indivisible phenomenon “D & E”:

The single components of the indivisible phenomenon known as “wire pendulum” are coupled by electromagnetic or gravitational interactions. The molecules of the sphere, the molecules of the wire, and the molecules of the fixed point adhere to one another through the electromagnetic interaction of their electron shells. And the sphere interacts with the Earth through gravitational attraction.

In contrast, there is *no* interaction⁵³ between the flying atoms D and E. The coupling between atoms D and E is not based on any interaction, but on *entanglement*. There is no such thing in classical physics. That is why Bohr quite rightly speaks here of a “feature of *individuality* completely foreign to classical physics” [39], and Schrödinger, with good reason, described entanglement not as “*one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought.” [32]

Entanglement also provides a clear criterion for determining when (only!) classical physics is suitable for describing a phenomenon, and when (only!) quantum theory is. If the phenomenon is merely a part of an entangled system (but not the entire entangled system), then only classical physics can be used to describe it. This is because in quantum theory, there are no state vectors for describing the parts of an entangled system. We saw this above when comparing the state vectors (6.3), (6.4), and (6.5).

Humans are always just parts of entangled systems because they must constantly breathe air. As a result, they become entangled with the air molecules, and the objects, which these molecules strike later, also become parts of the entangled overall system “humans and environment”. The same applies to all objects in our daily lives, and also to the wire pendulum shown in fig. 6.5. Quantum theory is applicable only to the entangled overall system, but that system is too large and too complex for a physical description. Therefore, in our physical analysis, we must limit ourselves to parts of the entangled overall system, and therefore only classical physics is applicable.

A description using quantum theory is only possible if uncontrolled entanglement with the environment is sufficiently suppressed

⁵³ The tiny, theoretically existing gravitational interaction between the two atoms plays absolutely no role.

through careful experimental setup, particularly by evacuating the apparatus.⁵⁴ If parts of such an isolated system are entangled with one another (such as the atoms D and E flying away after the decay of the molecule DE), then *only* a description using quantum theory is possible, because classical physics cannot describe entanglement.

In chapter 9 I will have quite a bit more to write on this topic.

6.5 Bell's Inequality

For nearly three decades, there was no real progress in the debate initiated by Einstein, Podolsky, and Rosen regarding the completeness or incompleteness of quantum theory. It was not until 1964 that things began to move again, thanks to an article [40] by John Bell (1928–1990) titled “On the Einstein Podolski Rosen Paradox”. Bell derived an inequality that applies to all theories assuming that alternative B is correct, i. e. that the magnetic moments of atoms along arbitrary spatial axes, and numerous other properties of quantum objects — which, according to quantum theory, are only generated by measurement and do not exist beforehand — actually already exist prior to measurement and are merely detected by the measurement.

To understand Bell's inequality, we must first translate the results of the measurements with Stern-Gerlach-magnets into numerical values r_{γ_D} and r_{γ_E} . We do this as follows:

$$\begin{array}{ll} \uparrow_{\gamma_D} \leftrightarrow r_{\gamma_D} = +1 & \downarrow_{\gamma_D} \leftrightarrow r_{\gamma_D} = -1 \\ \uparrow_{\gamma_E} \leftrightarrow r_{\gamma_E} = +1 & \downarrow_{\gamma_E} \leftrightarrow r_{\gamma_E} = -1 \end{array}$$

⁵⁴ Photons are a special case. They can travel kilometers through air and be deflected by mirrors, prisms, and lenses without becoming uncontrollably entangled with the environment. In this respect, however, photons are truly the absolute exception. Physics knows of no other type of objects which are even remotely as immune to entanglement with the environment as photons.

A measurement using the Stern-Gerlach-magnet can result in the atom being deflected toward the north pole (we denote this result by \uparrow or by $r = +1$), or in the atom being deflected toward the south pole (we denote this result by \downarrow or by $r = -1$). In addition, the subscript γ indicates the angle at which the magnet was set when the result was obtained; see fig. 6.3 on page 136.

Now let's tentatively assume that alternative B favored by EPR is correct: Although the magnetic moments of atoms D and E can be measured in the direction of only one spatial axis each with the Stern-Gerlach-magnets (and in this measurement the two results r_{γ_D} and r_{γ_E} are obtained), the magnetic moments must exist just as real in the direction of any other spatial axes. For, based on the correlation (6.7), we would e. g. be able to predict with probability $P = 1$ (i. e. with complete certainty) the result $r_{\gamma'_E}$ with $\gamma'_E \neq \gamma_E$ by measuring r_{γ_D} with $\gamma_D = \gamma'_E$. And what can be predicted with probability $P = 1$ that must — according to the EPR argument — actually exist in reality.

For our purpose, it suffices to assume that, in addition to the two magnetic moments measured with Stern-Gerlach-magnets set to γ_D and γ_E , two further magnetic moments *which could have been measured with Stern-Gerlach-magnets set to γ'_D and γ'_E but actually have not been measured*, exist just as really. Bell's derivation of his inequality is far too complicated for this book, but Asher Peres (1934–2005) later showed that there is a much simpler way [41]. Actually the simple assumption

A1_{Peres}: The quartet

$$(r_{\gamma_D}, r_{\gamma_E}, r_{\gamma'_D}, r_{\gamma'_E})$$

with two actually measured values and in addition two not measured values exists just as really as the doublet

$$(r_{\gamma_D}, r_{\gamma_E})$$

of the two actually measured values.

is already sufficient. Since every r is either $+1$ or -1 , 16 different

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
r_{γ_D}	+1	+1	+1	+1	+1	+1	+1	+1	-1	-1	-1	-1	-1	-1	-1	-1
r_{γ_E}	+1	+1	+1	+1	-1	-1	-1	-1	+1	+1	+1	+1	-1	-1	-1	-1
$r_{\gamma'_D}$	+1	+1	-1	-1	+1	+1	-1	-1	+1	+1	-1	-1	+1	+1	-1	-1
$r_{\gamma'_E}$	+1	-1	+1	-1	+1	-1	+1	-1	+1	-1	+1	-1	+1	-1	+1	-1
q	+2	+2	+2	-2	-2	-2	+2	-2	-2	+2	-2	-2	-2	+2	+2	+2

Tab. 6.1: The 16 quartets $(r_{\gamma_D}, r_{\gamma_E}, r_{\gamma'_D}, r_{\gamma'_E})$

quartets can be formed, which are listed in the 16 columns of table 6.1. The table is complete; besides these 16 there are no other quartets. Therefore, every time a DE molecule decays and the atoms D and E are analyzed using Stern-Gerlach-magnets, the result must be one of the 16 quartets from this table.

In the bottom line of the table, for each quartet the value of

$$q = r_{\gamma_D} \cdot r_{\gamma_E} + r_{\gamma_D} \cdot r_{\gamma'_E} + r_{\gamma'_D} \cdot r_{\gamma_E} - r_{\gamma'_D} \cdot r_{\gamma'_E} \quad (6.13)$$

is indicated. As the table is complete, there are no other values of q besides the 16 listed. Thus we know with certainty that in each measurement

$$q = +2 \quad \text{or} \quad q = -2. \quad (6.14)$$

We mark mean values by a bar. \bar{q} is the mean value of all values of q , which are found in a large number of runs of the experiment. Because of (6.14),

$$-2 \leq \bar{q} \leq +2 \quad (6.15)$$

$$\text{with } \bar{q} \stackrel{(6.13)}{=} \overline{r_{\gamma_D} \cdot r_{\gamma_E}} + \overline{r_{\gamma_D} \cdot r_{\gamma'_E}} + \overline{r_{\gamma'_D} \cdot r_{\gamma_E}} - \overline{r_{\gamma'_D} \cdot r_{\gamma'_E}}$$

must hold, no matter which result quartets are accidentally obtained.

The mean values, in turn, are closely related to the probabilities P of the various outcomes:

If $r_{\gamma_D} = +1$ and $r_{\gamma_E} = +1$, then $r_{\gamma_D} \cdot r_{\gamma_E} = +1$. The probability that this event occurs is

$$P(r_{\gamma_D} = +1, r_{\gamma_E} = +1) = P(\uparrow_{\gamma_D} \uparrow_{\gamma_E}).$$

If $r_{\gamma_D} = +1$ and $r_{\gamma_E} = -1$, then $r_{\gamma_D} \cdot r_{\gamma_E} = -1$. The probability that this event occurs is

$$P(r_{\gamma_D} = +1, r_{\gamma_E} = -1) = P(\uparrow_{\gamma_D} \downarrow_{\gamma_E}).$$

If $r_{\gamma_D} = -1$ and $r_{\gamma_E} = +1$, then $r_{\gamma_D} \cdot r_{\gamma_E} = -1$. The probability that this event occurs is

$$P(r_{\gamma_D} = -1, r_{\gamma_E} = +1) = P(\downarrow_{\gamma_D} \uparrow_{\gamma_E}).$$

If $r_{\gamma_D} = -1$ and $r_{\gamma_E} = -1$, then $r_{\gamma_D} \cdot r_{\gamma_E} = +1$. The probability that this event occurs is

$$P(r_{\gamma_D} = -1, r_{\gamma_E} = -1) = P(\downarrow_{\gamma_D} \downarrow_{\gamma_E}).$$

$$\begin{aligned} \text{Consequently } \overline{r_{\gamma_D} \cdot r_{\gamma_E}} &= +1 \cdot (P(\uparrow_{\gamma_D} \uparrow_{\gamma_E}) + P(\downarrow_{\gamma_D} \downarrow_{\gamma_E})) \\ &\quad - 1 \cdot (P(\uparrow_{\gamma_D} \downarrow_{\gamma_E}) + P(\downarrow_{\gamma_D} \uparrow_{\gamma_E})) \quad , \end{aligned}$$

and the inequality (6.15) can be written in the form

$$-2 \leq \bar{q} \leq +2 \tag{6.16}$$

$$\text{with } \bar{q} = \overline{r_{\gamma_D} \cdot r_{\gamma_E}} + \overline{r_{\gamma_D} \cdot r_{\gamma'_E}} + \overline{r_{\gamma'_D} \cdot r_{\gamma_E}} - \overline{r_{\gamma'_D} \cdot r_{\gamma'_E}}$$

$$\begin{aligned} \text{with } \overline{r_{\gamma_D} \cdot r_{\gamma_E}} &= P(\uparrow_{\gamma_D} \uparrow_{\gamma_E}) + P(\downarrow_{\gamma_D} \downarrow_{\gamma_E}) - \\ &\quad - P(\uparrow_{\gamma_D} \downarrow_{\gamma_E}) - P(\downarrow_{\gamma_D} \uparrow_{\gamma_E}) . \end{aligned}$$

(6.16) is Bell's inequality.

With the probabilities indicated above in (6.12), the correlation function becomes

$$\overline{r_{\gamma_D} \cdot r_{\gamma_E}} = \sin^2(\gamma_E/2 - \gamma_D/2) - \cos^2(\gamma_E/2 - \gamma_D/2) . \tag{6.17}$$

If \bar{q} is computed with this correlation function, Bell's inequality is violated with many angle settings. For example, with

$$\gamma_D = 45^\circ, \gamma_E = 90^\circ, \gamma'_D = 135^\circ, \gamma'_E = 0^\circ$$

we get

$$\bar{q} = -2\sqrt{2} \approx -2.83 \not\geq -2. \quad (6.18)$$

According to Bell's inequality (6.16), \bar{q} should never be less than -2 (and never greater than $+2$). Thus, with the correlation function (6.17) the inequality is violated. Here I must add, however, that the projection amplitudes (6.1), onto which the probabilities given in (6.12) are based, were in fact *not* obtained from series of measurements but calculated using quantum theory. In (6.1), I merely pointed out the possibility of measurement to spare readers the trouble of calculating it. The inconsistency (6.18) therefore merely proves that Bell's inequality is not compatible with quantum theory. This inconsistency does not yet prove, however, that the inequality is incompatible with the reality "out there".

What does the discrepancy (6.18) mean? If our derivation of Bell's inequality (6.16) contains no formal errors — and that is almost impossible in such a simple derivation — then there are only two possibilities:

- ① Either the correlation functions $\overline{r_{\gamma_D} \cdot r_{\gamma_E}} = (6.17)$, which have been computed based on quantum theory, are wrong. That would mean, then, that quantum theory is wrong.
- ② or assumption **A1**_{Peres}, onto which our derivation of Bell's inequality is based, is wrong, i. e. the assumption that not-measured results like e. g. $r_{\gamma'_D}$ and $r_{\gamma'_E}$ exist just as really as the measured results r_{γ_D} and r_{γ_E} .

In fact, the derivation of Bell's inequality is based not only on assumption $A1_{\text{Peres}}$. Bell later realized [42, penultimate section] that the following additional assumption is required:

A2_{Peres}: The settings of the measuring instruments (i. e. the angles to which the Stern-Gerlach-magnets resp. polarizers are adjusted), and the properties of the particles to be measured, are *not* predetermined by a common cause. (No “super-determinism”)

And Cramer [43, 44] eventually noted a third necessary assumption:

A3_{Peres}: The future result of a measurement does *not* affect the settings of the measuring instruments (i. e. the angles to which the Stern-Gerlach-magnets resp. polarizers are adjusted). (No “backward-causation”)

If possibility ② holds, then at least one of Peres' three basic assumptions $A1_{\text{Peres}}$, $A2_{\text{Peres}}$, $A3_{\text{Peres}}$ must be false. But which one? There are indeed some physicists who cast doubt on the seemingly almost self-evident assumptions $A2_{\text{Peres}}$ and $A3_{\text{Peres}}$. In line with the vast majority of physicists, however, I will assume in the following discussion that these two assumptions are correct. This means: If possibility ② holds, then I will conclude that assumption $A1_{\text{Peres}}$ (and thus also assumption B) is false.

Above all, however, it turned out that these questions can be discussed not only philosophically, but that it is possible to determine experimentally whether possibility ① or possibility ② holds true. To do this, one must measure the correlation function $\overline{r_{\gamma_D} \cdot r_{\gamma_E}} = (6.17)$ on suitable systems, and calculate the function $\overline{q} = (6.16)$ from the measurement results.

If there are detector settings $\gamma_D, \gamma'_D, \gamma_E, \gamma'_E$, for which $\overline{q} > 2$

or $\bar{q} < -2$ holds,⁵⁵ then possibility ② applies. Then⁵⁶ assumption **A1**_{Peres} — and thus, at the same time, assumption **B** — is false and experimentally refuted. Experiments proving that possibility ② actually holds true are presented in the following chapter.

⁵⁵ There's a logical stumbling block here: In assumption **A1**_{Peres}, r_{γ_D} and r_{γ_E} have been defined as really measured results, $r_{\gamma'_D}$ and $r_{\gamma'_E}$ as results of merely imagined but not actually carried-out measurements. How can $\bar{q} \stackrel{(6.16)}{=} \overline{r_{\gamma_D} \cdot r_{\gamma_E}} + \overline{r_{\gamma_D} \cdot r_{\gamma'_E}} + \overline{r_{\gamma'_D} \cdot r_{\gamma_E}} - \overline{r_{\gamma'_D} \cdot r_{\gamma'_E}}$ be determined experimentally, if no measurements are carried out with γ'_D and γ'_E ? Well, one performs a series of measurements with γ_D, γ_E , one series of measurements with γ_D, γ'_E , one series of measurements with γ'_D, γ_E , and one series of measurements with γ'_D, γ'_E , and determines from these four series of measurements the four correlation functions $\overline{r_{\gamma_D} \cdot r_{\gamma_E}}$ in (6.16). Thereby one must assume, of course, that the — according to **A1** merely hypothetical, but according to EPR entirely real and with probability $P = 1$ determined — results not suddenly change when they are actually measured. This is a very plausible assumption, however, one to that EPR would certainly have agreed without reservation.

⁵⁶ assuming that assumptions **A2**_{Peres} (no superdeterminism) and **A3**_{Peres} (no backward-causation) both are correct

This page is intentionally empty.

7 Experiments on Bell's Inequality

Der experimentelle Beweis, dass die Bell'sche Ungleichung tatsächlich verletzt wird, wurde seit den achtziger Jahren in zahlreichen Experimenten erbracht. Allerdings kennt man kein instabiles DE-Molekül mit magnetischem Moment $\uparrow\downarrow$ Null, das in die Atome D und E mit magnetischen Momenten \uparrow oder \downarrow zerfällt. Bohms in Abb. 6.3 auf Seite 136 skizziertes Experiment ist lediglich ein sogenanntes *Gedankenexperiment*. Es ist nützlich, um das Argument von EPR zu erklären, und es zeigt wie ein Experiment zum Test von Bells Ungleichung im Prinzip gestaltet werden könnte.

Beweiskraft haben aber natürlich nur Experimente, die tatsächlich im Labor durchgeführt wurden. Im Folgenden wird über mehrere derartige Experimente detailliert berichtet. Ich beschreibe diese Experimente deshalb so genau, damit erkennbar wird dass es sich um so solide, vertrauenswürdige, und zuverlässige Beweise handelt, wie man sie sich überhaupt nur wünschen und vorstellen kann. Die Experimente zeigen, dass die Annahme **B** wirklich definitiv und endgültig widerlegt ist.⁵⁶

Alle Experimente laufen nach dem gleichen einfachen Schema ab:

- * Erster Schritt: Ein Quantensystem, das aus zwei Teilen besteht, wird in einem verschränkten Zustand präpariert. Sein Zustandsvektor hat also die Form (6.3). Diese Art von Zustandsvektoren kann nicht als Produkt von Zustandsvektoren der Teilsysteme geschrieben werden, d. h. die Quantentheorie ordnet nur dem verschränkten Gesamtsystem einen Zustands-

vektor zu, aber nicht den Teilsystemen.

- * Zweiter Schritt: Eine bestimmte Eigenschaft wird an beiden Teilsystemen gemessen. Es stellt sich heraus, dass die Ergebnisse der Messungen an den beiden Teilsystemen korreliert sind, und zwar genau so wie durch die quantentheoretisch berechnete Korrelationsfunktion $\overline{r_{\gamma_D} \cdot r_{\gamma_E}} = (6.17)$ vorhergesagt.
- * Dritter Schritt: Die gemessenen Korrelationen $r_{\gamma_D} \cdot r_{\gamma_E}$ verletzen bei bestimmten Einstellungen der Messgeräte deutlich die Bell'sche Ungleichung (6.16). In Abschnitt 7.5 werden anschließend die Schlussfolgerungen aus diesem Resultat gezogen.

Ich habe mir die allergrößte Mühe gegeben die Experimente so klar wie möglich darzustellen. Einige technische Details sind aber so schwierig, dass ein großer Teil der Nicht-Physiker, an die sich dies Buch doch eigentlich in erster Linie richtet, hier aus der Kurve fliegen könnte. Diese besonders schwierigen Punkte habe ich deshalb in den Anhang verbannt, und empfehle allen Lesern – falls sie keine Physiker sind – zumindest beim ersten Lesen diese Anhänge zu ignorieren, um den „roten Faden“ nicht zu verlieren.

7.1 Verschränkte Ca-Lumineszenz-Photonen

In Abbildung 7.1 auf der nächsten Seite ist der Aufbau eines Experiments skizziert, das 1981 von Aspect, Grangier, und Roger [45, 46] durchgeführt wurde. Der türkise Punkt soll einen Strahl von Calcium-Atomen symbolisieren, die sich in einer evakuierten Kammer parallel zur y -Achse – also senkrecht zur Papierebene – bewegen.

Der Atomstrahl hat einen Durchmesser von etwa 1 mm, und enthält pro mm^3 etwa $3 \cdot 10^7$ Ca-Atome. Der Atomstrahl wird parallel und antiparallel zur x -Achse mit zwei Laserstrahlen angeleuchtet. Die Wellenlänge des einen Laserstrahls ist 406,7 nm, die Wellenlänge des anderen Laserstrahls ist 581 nm. Beide Laserstrahlen sind

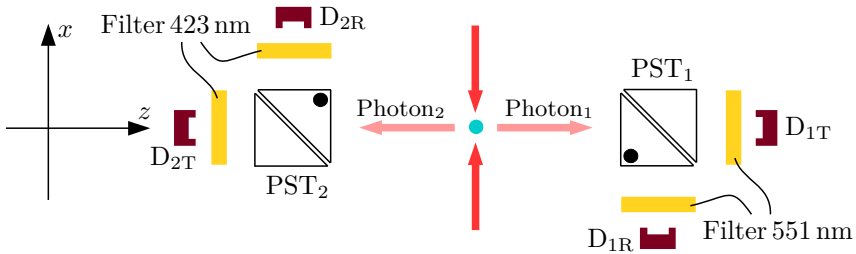


Fig. 7.1 : Das Experiment von Aspect, Grangier, und Roger [46]

parallel zur y -Achse (also senkrecht zur Papierebene der Skizze 7.1) polarisiert.

Aus dem Term-Schema 7.2 auf der nächsten Seite kann man ablesen, was bei der Wechselwirkung der Calcium-Atome mit den beiden Laserstrahlen vor sich geht. Die drei waagerechten Striche symbolisieren die Energie des Calcium-Atoms in verschiedenen Zuständen. Je höher der Strich, desto höher die Energie. Der unterste Strich symbolisiert die Energie des Atoms im Grundzustand. Durch die gleichzeitige Absorption von zwei Photonen,⁵⁷ eines mit Wellenlänge 406,7 nm und eines mit Wellenlänge 581 nm, wird es angeregt in einen Zustand, der durch den obersten Strich symbolisiert wird. Nach einiger Zeit emittiert das Atom ein Photon mit Wellenlänge 551,3 nm, so dass es nur noch die Energie hat, die durch den mittleren Strich symbolisiert wird. Mit einer Halbwertszeit von 5 ns emittiert das Atom ein weiteres Photon mit Wellenlänge 422,7 nm und gelangt dadurch wieder in den Grundzustand.

⁵⁷ Bei der Besprechung von Lenards Experimenten zum Lichtelektrischen Effekt in Abschnitt 3.1 wurde gesagt, dass die gleichzeitige Absorption von zwei Photonen an der gleichen Stelle extrem unwahrscheinlich ist. Aber Lenards Lichtquellen waren wirklich nur kläglich trübe Funzeln im Vergleich zu der geballten Laser-Power, die den Experimentatoren 1981 zur Verfügung stand. Bei dieser hohen Intensität des anregenden Lichts kommt die gleichzeitige Absorption von zwei Photonen häufig vor.

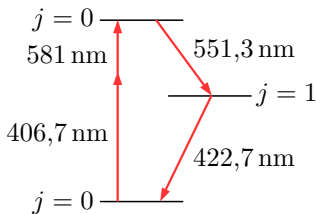


Abb. 7.2: Eine Lumineszenz-Kaskade des Calcium-Atoms

Neben den Strichen sind die Drehimpuls-Quantenzahlen des Atoms in diesen drei Zuständen angegeben. **Drehimpuls** kann weder aus dem Nichts auftauchen noch ins Nichts verschwinden. Wenn sich der Drehimpuls des Atoms bei den beiden Übergängen der Lumineszenzkaskade ändert, dann müssen die emittierten Lumineszenz-Photonen den fehlenden Drehimpuls mitgenommen haben.

Die Erhaltung des Drehimpulses erzwingt die folgende Form der Zustandsfunktion⁵⁸ der beiden Lumineszenz-Photonen:

$$|\text{Photon}_1 \& \text{Photon}_2\rangle \stackrel{\text{(A.10c)}}{=} \sqrt{\frac{1}{2}} \left(|L_\gamma\rangle_1 |L_\gamma\rangle_2 + |L_{\gamma+90}\rangle_1 |L_{\gamma+90}\rangle_2 \right) \\ \text{mit beliebigem } \gamma \quad (7.1)$$

$|L_\gamma\rangle$ ist der Zustandsvektor eines linear polarisierten⁵⁹ Photons, das sich entlang der z -Achse des Koordinatensystems bewegt, und dessen Polarisations Ebene um den Winkel γ gegen die y -Achse gedreht ist, siehe Abb. 7.3 auf der nächsten Seite. Der Winkel γ kann jeden beliebigen Wert $0^\circ \leq \gamma \leq 180^\circ$ haben.

⁵⁸ Mathematisch besonders interessierte Leser können die Begründung von (7.1) in Anhang A.2 nachlesen. Alle anderen sollten einfach glauben, dass der Zustandsvektor (7.1) korrekt ist.

⁵⁹ Lineare Polarisation ist die Art von Polarisation, die Photonen mithilfe polarisierender Strahlteiler aufgeprägt wird, wie in Abschnitt 2.2 beschrieben. Ein Photon, das von einem um den Winkel γ gedrehten polarisierenden Strahlteiler (siehe Abb. 2.9 auf Seite 36) *transmittiert* (nicht reflektiert) wurde, hat die Polarisation L_γ . Es gibt auch noch andere Arten der Polarisation, mit denen wir uns aber in diesem Buch nicht zu befassen brauchen. Wer es trotzdem wissen will, sollte Anhang A.2 lesen.

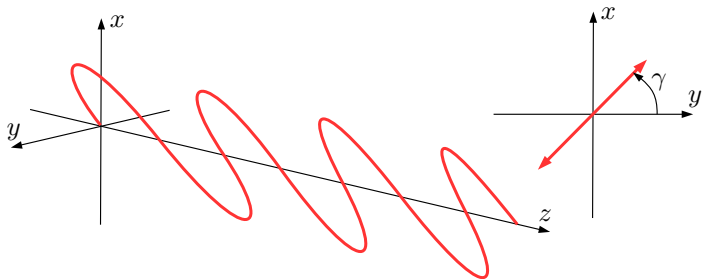


Fig. 7.3: Der Polarisationswinkel γ der Lumineszenz-Photonen

Der Zustandsvektor (7.1) beschreibt die Polarisation des verschränkten Gesamtsystems $\text{Photon}_1 \& \text{Photon}_2$, er beschreibt aber nicht die Polarisation der einzelnen Photonen. Um das klar zu sehen, betrachten wir zum Vergleich diese Zustandsvektoren:

$$|\text{Photon}\rangle_1 = u_1 |L_\gamma\rangle_1 + v_1 |L_{\gamma+90}\rangle_1 \quad \text{mit } |u_1|^2 + |v_1|^2 = 1 \quad (7.2a)$$

$$|\text{Photon}\rangle_2 = u_2 |L_\gamma\rangle_2 + v_2 |L_{\gamma+90}\rangle_2 \quad \text{mit } |u_2|^2 + |v_2|^2 = 1 \quad (7.2b)$$

Dies sind die allgemeinsten Zustandsvektoren die man für einzelne Photonen aufschreiben kann. Wenn jedes der beiden Photonen als eigenständiges Quantenobjekt existiert, dann ist der Zustandsvektor des Gesamtsystems gleich dem Produkt dieser Vektoren:

$$\begin{aligned} |\text{Photon}\rangle_1 |\text{Photon}\rangle_2 &= \\ &= \left(u_1 |L_\gamma\rangle_1 + v_1 |L_{\gamma+90}\rangle_1 \right) \left(u_2 |L_\gamma\rangle_2 + v_2 |L_{\gamma+90}\rangle_2 \right) = \\ &= u_1 u_2 |L_\gamma\rangle_1 |L_\gamma\rangle_2 + u_1 v_2 |L_\gamma\rangle_1 |L_{\gamma+90}\rangle_2 + \\ &\quad + v_1 u_2 |L_{\gamma+90}\rangle_1 |L_\gamma\rangle_2 + v_1 v_2 |L_{\gamma+90}\rangle_1 |L_{\gamma+90}\rangle_2 \quad (7.2c) \\ &\quad \text{mit } |u_1|^2 + |v_1|^2 = 1 \text{ und } |u_2|^2 + |v_2|^2 = 1 \end{aligned}$$

Der verschränkte Zustandsvektor (7.1) unterscheidet sich fundamental von (7.2c). Denn wenn man (7.2c) auf die Form (7.1) zurückführen wollte, dann müsste

$$u_1 v_2 = 0 = v_1 u_2 \quad \text{und} \quad u_1 u_2 \neq 0 \neq v_1 v_2 \quad (7.2d)$$

sein, aber diese Bedingungen können unmöglich gleichzeitig erfüllt werden.

Der verschränkte Zustandsvektor (7.1) kann also nicht als Produkt der Zustandsvektoren (7.2) einzelner Photonen geschrieben werden. Die Quantentheorie ordnet den Teilen eines verschränkten Systems *überhaupt keinen* Zustandsvektor zu. In dem verschränkten Zustand (7.1) sind die Zustandsvektoren $|L_\gamma\rangle_1$, $|L_{\gamma+90}\rangle_1$, $|L_\gamma\rangle_2$, $|L_{\gamma+90}\rangle_2$ der beiden einzelnen Photonen zwar enthalten, man kann (7.1) aber nicht faktorisieren und in die Form (7.2c) bringen, in der für jedes Photon ein Zustandsvektor definiert ist. Im Formalismus der Quantentheorie haben die beiden Lumineszenz-Photonen keine eigenständige Existenz, sondern existieren nur als Bestandteile des Gesamtsystems „zwei Lumineszenzphotonen“.

Um die Korrelation der Polarisation der beiden Lumineszenz-Photonen zu prüfen, wurden im Experiment von Aspect et al. polarisierende Strahlteiler eingesetzt, deren Funktionsweise in Abb. 2.8 auf Seite 35 dargestellt wird. Photon₁ bewegt sich in Richtung der positiven z -Achse zum polarisierenden Strahlteiler PST₁, und Photon₂ bewegt sich in Richtung der negativen z -Achse zum polarisierenden Strahlteiler PST₂, siehe Abb. 7.1 auf Seite 160.

Zwischen PST₁ und den Detektoren D_{1T} und D_{1R} befindet sich je ein Filter, der nur Photonen mit einer Wellenlänge von etwa 551 nm durchlässt. Zwischen PST₂ und den Detektoren D_{2T} und D_{2R} befindet sich je ein Filter, der nur Photonen mit einer Wellenlänge von etwa 423 nm durchlässt. Diese Filter sind erforderlich, weil sonst direktes Streulicht von den anregenden Laserstrahlen das schwache Lumineszenzlicht überdecken würde.

Wenn Photon₁ eine Wellenlänge von etwa 551 nm hat, dann wird es von Detektor D_{1T} detektiert falls es vom polarisierenden Strahlteiler transmittiert wird, bzw. von Detektor D_{1R} detektiert falls es vom polarisierenden Strahlteiler reflektiert wird. Wenn

Photon₂ eine Wellenlänge von etwa 423 nm hat, dann wird es von Detektor D_{2T} detektiert falls es vom polarisierenden Strahlteiler transmittiert wird, bzw. von Detektor D_{2R} detektiert falls es vom polarisierenden Strahlteiler reflektiert wird.

Jeder der vier Detektoren zählte im Experiment von Aspect, Grangier, und Roger etwa 10^4 Photonen pro Sekunde. Eine Korrelation der Polarisation ist aber nur zwischen den jeweils zwei Photonen der Lumineszenz-Kaskade des gleichen Atoms zu erwarten, nicht zwischen den Lumineszenz-Photonen verschiedener Atome. Deshalb wurden die Ergebnisse der vier Zähler zeitaufgelöst analysiert: Wenn einer der beiden Zähler D_{1T} oder D_{1R} ein Photon registrierte, dann wurde geprüft ob

- * innerhalb von 20 ns der jeweils andere der beiden Zähler D_{1T} oder D_{1R} *kein* Photon registrierte,
- * im gleichen 20 ns-Zeitraum einer der beiden Zähler D_{2T} oder D_{2R} ein Photon registrierte,
- * der jeweils andere der beiden Zähler D_{2T} oder D_{2R} im gleichen 20 ns-Zeitraum *kein* Photon registrierte.

Wenn alle drei Bedingungen erfüllt waren, dann wurden die beiden Photonen als Lumineszenzphotonen des gleichen Atoms interpretiert, und das Ereignis wurde gespeichert. Wenn mindestens eine Bedingung nicht erfüllt war, wurde das Ereignis ignoriert. Abhängig von der Stellung der Polarisatoren wurden mit diesen Kriterien bis zu 50 gültige Ereignisse pro Sekunde registriert. Weil im Term-schema 7.2 die Halbwertszeit des mittleren Niveaus 5 ns beträgt, wird das zweite Lumineszenz-Photon der Kaskade tatsächlich nur selten außerhalb des 20 ns langen Zeitfensters liegen. Aber das umgekehrte gilt nicht: Dass ein Ereignis alle drei Bedingungen erfüllt beweist noch nicht zwingend, dass beide Photonen wirklich aus der Lumineszenz-Kaskade des gleichen Atoms stammen. Aspect et al. schätzten vielmehr (und konnten das auch mit geeigneten Tests

belegen), dass ziemlich stabil etwa 10 scheinbar gültige Ereignisse pro Sekunde registriert wurden, bei denen die beiden detektierten Photonen tatsächlich nicht aus der gleichen Kaskade stammten sondern nur zufällig ins gleiche Zeitfenster gerutscht waren. Deshalb zogen sie von allen ihren Ergebnissen 10 Ereignisse pro Sekunde ab. Damit blieben letztlich noch – je nach Stellung der Polarisatoren – bis zu 40 gültige Ereignisse pro Sekunde übrig.

Die beiden polarisierenden Strahlteiler konnten unabhängig voneinander mitsamt ihren nachgeschalteten Filtern und Detektoren um die z -Achse gedreht werden. In der rechten Skizze von Abb. 7.4 wird gezeigt, wie das System der Koordinaten a und b definiert wird, das sich mit den Strahlteilern um die z -Achse dreht. Aus den beiden anderen Skizzen kann man ablesen, wie die Drehwinkel γ_1 und γ_2 der beiden polarisierenden Strahlteiler definiert werden, nämlich als Winkel zwischen der positiven y -Achse und der positiven b_1 -Achse bzw. der negativen b_2 -Achse. In der mittleren Skizze blickt man in Richtung der positiven z -Achse, also in Bewegungsrichtung von Photon_1 , auf den Strahlteiler PST_1 . In der linken Skizze blickt man in Richtung der negativen z -Achse, also in Bewegungsrichtung von Photon_2 , auf den Strahlteiler PST_2 .

Wir definieren die Einheitsvektoren $|a_1\rangle$, $|b_1\rangle$, $|a_2\rangle$, $|b_2\rangle$ parallel zu den jeweiligen Koordinatenachsen a und b . Die vier Eigenvektoren der Messapparatur (sprich der beiden polarisierenden Strahlteiler

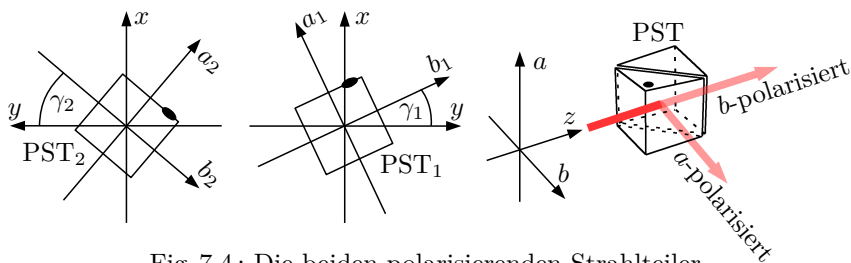


Fig. 7.4: Die beiden polarisierenden Strahlteiler

mit den nachgeschalteten Detektoren) sind also

$$|a_1\rangle |a_2\rangle \quad , \quad |a_1\rangle |b_2\rangle \quad , \quad |b_1\rangle |a_2\rangle \quad , \quad |b_1\rangle |b_2\rangle . \quad (7.3)$$

Wir nennen

W_{RR} die Wahrscheinlichkeit dafür dass beide Photonen reflektiert werden (also nach der Messung in a -Richtung des jeweiligen Strahlteilers polarisiert sind und durch den Zustandsvektor $|a_1\rangle |a_2\rangle$ beschrieben werden),

W_{RT} die Wahrscheinlichkeit dafür dass Photon₁ reflektiert und Photon₂ transmittiert wird (die Photonen also nach der Messung in a_1 -Richtung bzw. in b_2 -Richtung polarisiert sind und durch den Zustandsvektor $|a_1\rangle |b_2\rangle$ beschrieben werden),

W_{TR} die Wahrscheinlichkeit dafür dass Photon₁ transmittiert und Photon₂ reflektiert wird (die Photonen also nach der Messung in b_1 -Richtung bzw. in a_2 -Richtung polarisiert sind und durch den Zustandsvektor $|b_1\rangle |a_2\rangle$ beschrieben werden), und

W_{TT} die Wahrscheinlichkeit dafür dass beide Photonen von ihrem jeweiligen Strahlteiler transmittiert werden (also nach der Messung in b -Richtung des jeweiligen Strahlteilers polarisiert sind und durch den Zustandsvektor $|b_1\rangle |b_2\rangle$ beschrieben werden).

Um diese Wahrscheinlichkeiten zu berechnen, projizieren wir wie gewohnt zunächst den Zustandsvektor (7.1) des Photonenpaars vor der Messung auf die vier Eigenvektoren (7.3) der Messapparatur:

$$\begin{aligned} |\text{Photon}_1 \& \text{Photon}_2\rangle \stackrel{(7.1)}{=} \sqrt{\frac{1}{2}} \left[\right. \\ & |a_1\rangle |a_2\rangle \left(\langle a_1 || L_\gamma \rangle_1 \langle a_2 || L_\gamma \rangle_2 + \langle a_1 || L_{\gamma+90} \rangle_1 \langle a_2 || L_{\gamma+90} \rangle_2 \right) + \\ & + |a_1\rangle |b_2\rangle \left(\langle a_1 || L_\gamma \rangle_1 \langle b_2 || L_\gamma \rangle_2 + \langle a_1 || L_{\gamma+90} \rangle_1 \langle b_2 || L_{\gamma+90} \rangle_2 \right) + \\ & + |b_1\rangle |a_2\rangle \left(\langle b_1 || L_\gamma \rangle_1 \langle a_2 || L_\gamma \rangle_2 + \langle b_1 || L_{\gamma+90} \rangle_1 \langle a_2 || L_{\gamma+90} \rangle_2 \right) + \\ & \left. + |b_1\rangle |b_2\rangle \left(\langle b_1 || L_\gamma \rangle_1 \langle b_2 || L_\gamma \rangle_2 + \langle b_1 || L_{\gamma+90} \rangle_1 \langle b_2 || L_{\gamma+90} \rangle_2 \right) \right] \end{aligned}$$

Nach der Born'schen Regel (5.9) ist die Wahrscheinlichkeit eines bestimmten Messergebnisses gleich dem Betragsquadrat der Projektionsamplitude auf den jeweiligen Eigenvektor der Messapparatur:

$$W_{RR} = \frac{1}{2} \left| \langle a_1 || L_\gamma \rangle_1 \langle a_2 || L_\gamma \rangle_2 + \langle a_1 || L_{\gamma+90} \rangle_1 \langle a_2 || L_{\gamma+90} \rangle_2 \right|^2 \quad (7.4a)$$

$$W_{RT} = \frac{1}{2} \left| \langle a_1 || L_\gamma \rangle_1 \langle b_2 || L_\gamma \rangle_2 + \langle a_1 || L_{\gamma+90} \rangle_1 \langle b_2 || L_{\gamma+90} \rangle_2 \right|^2 \quad (7.4b)$$

$$W_{TR} = \frac{1}{2} \left| \langle b_1 || L_\gamma \rangle_1 \langle a_2 || L_\gamma \rangle_2 + \langle b_1 || L_{\gamma+90} \rangle_1 \langle a_2 || L_{\gamma+90} \rangle_2 \right|^2 \quad (7.4c)$$

$$W_{TT} = \frac{1}{2} \left| \langle b_1 || L_\gamma \rangle_1 \langle b_2 || L_\gamma \rangle_2 + \langle b_1 || L_{\gamma+90} \rangle_1 \langle b_2 || L_{\gamma+90} \rangle_2 \right|^2 \quad (7.4d)$$

Die in Anhang A.3 detailliert erklärte Berechnung ergibt

$$W_{TT} \stackrel{(A.17b)}{=} W_{RR} \stackrel{(A.19)}{=} \frac{1}{2} \cos^2(\gamma_1 - \gamma_2) \quad (7.5a)$$

$$W_{RT} \stackrel{(A.17a)}{=} W_{TR} \stackrel{(A.17d)}{=} \frac{1}{2} - \frac{1}{2} \cos^2(\gamma_1 - \gamma_2) . \quad (7.5b)$$

Bemerkenswert ist, dass erstens der Winkel γ , der in der Zustandsfunktion $|\text{Photon}_1 \& \text{Photon}_2\rangle = (7.1)$ enthalten war, in (7.5) überhaupt nicht mehr auftaucht, und dass zweitens (7.5) nur von der Differenz $\gamma_2 - \gamma_1$ abhängt, aber nicht von den absoluten Werten dieser Winkel.

Mit (7.5) können wir die Korrelationsfunktion berechnen:

$$\begin{aligned} \overline{r_{\gamma_1} \cdot r_{\gamma_2}} &\stackrel{(6.16)}{=} W_{TT} + W_{RR} - W_{TR} - W_{RT} = \\ &= 2 \cos^2(\gamma_2 - \gamma_1) - 1 \end{aligned} \quad (7.6)$$

In der Korrelationsfunktion kommt das Quadrat \cos^2 der Kosinusfunktion vor, die in Abb. 5.4 auf Seite 109 grafisch dargestellt wurde. Aus dieser Grafik kann man mithilfe der roten Linie für

jede beliebige Winkeldifferenz $\gamma_2 - \gamma_1$ den Wert von $\cos^2(\gamma_2 - \gamma_1)$ ablesen.⁶⁰

Jetzt ist noch ein Korrekturfaktor erforderlich, und zwar aus folgendem Grund: Wir haben bisher angenommen, dass die beiden Lumineszenz-Photonen in *genau* entgegengesetzte Richtung emittiert werden. Vollkommene Genauigkeit ist bei keinem Experiment möglich, und der Winkel zwischen den Flugrichtungen der beiden Photonen kann auch geringfügig von 180° abweichen. Das hat eine geringfügig schwächere Korrelation der Polarisationsrichtungen zur Folge. Aspect et al. berechneten dafür den folgenden Korrekturfaktor:

$$K = 0,984 \quad (7.7)$$

Damit ergibt sich die theoretisch berechnete Korrelationsfunktion

$$K \cdot \overline{r_{\gamma_1} \cdot r_{\gamma_2}} \stackrel{(7.6)}{=} 0,984 \cdot \left(2 \cos^2(\gamma_2 - \gamma_1) - 1 \right). \quad (7.8)$$

Diese Funktion ist als blaue Kurve im Diagramm 7.5 auf der nächsten Seite eingetragen.

Um die Korrelationsfunktion experimentell zu überprüfen, zählten Aspect et al. bei verschiedenen Einstellungen der Winkel γ_1 und γ_2

die Zahl N_{TT} gültiger Ereignisse pro 100 Sekunden, bei denen die Detektoren D_{1T} und D_{2T} ansprachen,

die Zahl N_{RR} gültiger Ereignisse pro 100 Sekunden, bei denen die Detektoren D_{1R} und D_{2R} ansprachen,

die Zahl N_{TR} gültiger Ereignisse pro 100 Sekunden, bei denen die Detektoren D_{1T} und D_{2R} ansprachen,

die Zahl N_{RT} gültiger Ereignisse pro 100 Sekunden, bei denen die Detektoren D_{1R} und D_{2T} ansprachen.

Aus diesen Ergebnissen ergaben sich die Wahrscheinlichkeiten

⁶⁰ Wesentlich einfacher geht das mit jedem halbwegs brauchbaren Taschenrechner.

$$W_{TT} = \frac{N_{TT}}{N_{TT} + N_{RR} + N_{TR} + N_{RT}} \quad (7.9a)$$

$$W_{RR} = \frac{N_{RR}}{N_{TT} + N_{RR} + N_{TR} + N_{RT}} \quad (7.9b)$$

$$W_{TR} = \frac{N_{TR}}{N_{TT} + N_{RR} + N_{TR} + N_{RT}} \quad (7.9c)$$

$$W_{RT} = \frac{N_{RT}}{N_{TT} + N_{RR} + N_{TR} + N_{RT}} \quad (7.9d)$$

und die Korrelationsfunktion

$$\left(\overline{r_{\gamma_1} \cdot r_{\gamma_2}} \right)_{\text{experimentell}} \stackrel{(6.16)}{=} \left(W_{TT} + W_{RR} - W_{TR} - W_{RT} \right)_{\text{experimentell}}$$

Wie erwartet stellte sich heraus, dass die Korrelation $\overline{r_{\gamma_1} \cdot r_{\gamma_2}}$ nur von der Differenz $\gamma_2 - \gamma_1$ abhängt (z. B. war $\overline{r_{30^\circ} \cdot r_{75^\circ}} = \overline{r_{0^\circ} \cdot r_{45^\circ}}$, usw.). Die Ergebnisse für die Einstellungen $\gamma_1 - \gamma_2$ gleich 0° , $22,5^\circ$, 30° , 45° , 60° , $67,5^\circ$, und 90° sind im Diagramm 7.5 als rote Punkte eingetragen. Theorie (mit dem Korrekturfaktor $K \stackrel{(7.7)}{=} 0,984$) und Experiment stimmen offensichtlich sehr gut überein.

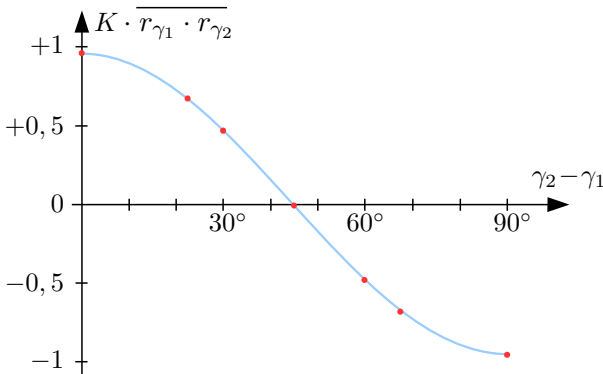


Fig. 7.5: Die Korrelation $F_K(\gamma_1, \gamma_2)$

Aspect et al. überprüften mit ihren Messdaten auch die in (6.16) beschriebene Bell'sche Ungleichung

$$-2 \leq \bar{q} \leq +2 \quad (7.10)$$

$$\text{mit } \bar{q} = \overline{r_{\gamma_1} \cdot r_{\gamma_2}} + \overline{r_{\gamma_1} \cdot r_{\gamma_2'}} + \overline{r_{\gamma_1'} \cdot r_{\gamma_2}} - \overline{r_{\gamma_1'} \cdot r_{\gamma_2'}} ,$$

die mit beliebigen Einstellungen γ_1 und γ_1' von PST₁ und beliebigen Einstellungen γ_2 und γ_2' von PST₂ erfüllt sein muss, falls die Annahme B richtig sein sollte.

Wenn dagegen die Quantentheorie richtig ist, dann wird die Ungleichung (7.10) bei zahlreichen Einstellungen der Strahlteiler verletzt, und zwar besonders stark bei den Winkelkombinationen

$$\begin{aligned} \gamma_2 - \gamma_1 = \gamma_1' - \gamma_2 = \gamma_2' - \gamma_1' &= 22,5^\circ \\ \gamma_2' - \gamma_1 &= 67,5^\circ . \end{aligned}$$

Bei diesen Winkeln berechnet man mit der Quantentheorie den Wert

$$K \cdot \bar{q}_{\text{Quantentheorie}} = K \cdot 2\sqrt{2} = 2,70 , \quad (7.11a)$$

wobei $K \stackrel{(7.7)}{=} 0,984$ der Korrekturfaktor dafür ist, dass die Lumineszenzphotonen sich nicht völlig exakt entlang der z -Achse bewegen. Aus den Messwerten ergab sich bei den gleichen Winkeln

$$\bar{q}_{\text{Experiment}} = 2,697 \pm 0,015 . \quad (7.11b)$$

Dies Experiment verletzt die Bell'sche Ungleichung (7.10) so deutlich, dass kein Raum für Alternative B bleibt. Das Experiment beweist, dass die Polarisation der Photonen nicht bereits bei der Emission festgelegt wird, sondern erst durch die Wechselwirkung mit den Strahlteilern und Detektoren erschaffen wird. Bei der Emission wird lediglich festgelegt, dass beide Photonen in der gleichen Richtung linear polarisiert sein werden, wenn (erst dann, vorher nicht!) ihre lineare Polarisation durch eine Messung in der Zukunft erschaffen werden sollte.

7.2 Verschränkte Be^+ Ionen

Unter allen in diesem Buch berichteten Experimenten ist dasjenige mit Beryllium-Ionen, das ich in diesem Abschnitt beschreibe, das bei weitem komplizierteste. Eigentlich ist es zu schwierig für dieses Buch, aber ich wollte unbedingt einen Test der Bell'schen Ungleichung mit materiellen Teilchen dabei haben. Sonst könnte der falsche Eindruck entstehen, dass nur Photonen die Bell'sche Ungleichung verletzen. Auch hier habe ich besonders schwierige technische Details in den Anhang verschoben, und wiederhole nochmals die Empfehlung, die Anhänge beim ersten Lesen zunächst zu ignorieren, um den „roten Faden“ nicht zu verlieren. Überhaupt sollten Leser, die mit dem mathematischen Formalismus nicht so vertraut sind, sich nicht scheuen diesen Abschnitt nur „diagonal“ zu lesen, statt bei den (zahlreichen!) sehr schwierigen Details stecken zu bleiben. Die danach folgenden Abschnitte sind dann wieder wesentlich leichter.

Im Jahr 2000 präparierten Rowe et al. [47] Paare von Beryllium-Ionen im verschränkten Zustand

$$|\text{Be}^+ \& \text{Be}^+\rangle = \sqrt{\frac{1}{2}} \left(|\downarrow\rangle_1 |\downarrow\rangle_2 - |\uparrow\rangle_1 |\uparrow\rangle_2 \right). \quad (7.12)$$

In Bohms Gedankenexperiment befinden sich die Atome D und E im Zustand (6.3), bei dem die magnetischen Momente in Richtung beliebiger Raumachsen – wenn sie durch eine Messung erschaffen werden – stets entgegengesetzt sind. Dagegen sind die magnetischen Momente der beiden Beryllium-Ionen im Zustand (7.12) – wenn sie durch eine Messung erschaffen werden – stets gleich gerichtet. Das ändert nichts daran, dass (7.12) ein verschränkter Zustand ist, denn er kann nicht als Produkt der Zustände

$$|\text{Be}^+\rangle_1 = q|\uparrow\rangle_1 + u|\downarrow\rangle_1 \quad \text{mit } |q|^2 + |u|^2 = 1 \quad (7.13a)$$

$$|\text{Be}^+\rangle_2 = v|\uparrow\rangle_2 + w|\downarrow\rangle_2 \quad \text{mit } |v|^2 + |w|^2 = 1 \quad (7.13b)$$

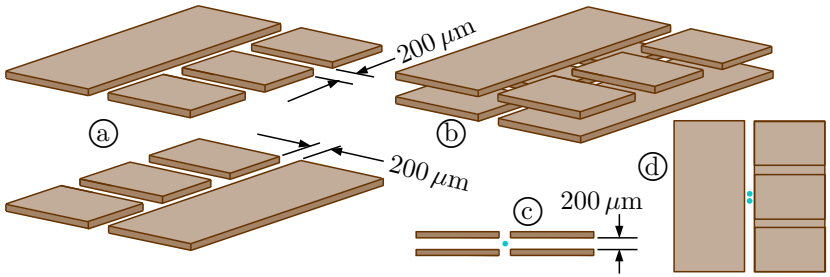


Fig. 7.6: Die Ionen-Falle

der einzelnen Ionen geschrieben werden, egal welche Werte man für die Amplituden q, u, v, w wählt.

$|\uparrow\rangle$ und $|\downarrow\rangle$ sind im Fall des Beryllium-Ions abkürzende Schreibweisen. Für das Verständnis des Experiments ist es unwichtig, um welche Zustände es sich im Detail handelt. Wer es trotzdem genau wissen möchte, kann es in Anhang A.4 nachlesen.

Um zwei Be^+ Ionen in den verschränkten Zustand (7.12) zu bringen, gingen Rowe et al. folgendermaßen vor: In einer evakuierten Kammer befand sich die elektrostatische Falle, die in Abb. 7.6 skizziert ist. In 7.6(b) schaut man von schräg oben, in 7.6(c) von der Stirnseite, und in 7.6(d) senkrecht von oben auf die Falle. In 7.6(a) sind zwecks besserer Sichtbarkeit die oberen und unteren Elektroden einzeln gezeichnet.

Die türkisen Punkte in 7.6(d) sollen die Be^+ Ionen symbolisieren. Gleichnamige Ladungen stoßen sich ab. Wenn man eine positive Spannung an die Elektroden legt, dann stoßen sie die positiv geladenen Ionen ab, so dass die Ionen im Inneren der Falle gefangen sind. (An die vier kleinen Elektroden in den Ecken wird eine etwas höhere positive Spannung angelegt als an die vier anderen Elektroden, damit die Ionen sich nicht durch den länglichen Kanal hinaus schleichen können.)

Wie bringt man die Ionen in die Falle hinein? Man schießt

zunächst einen Strahl von Ionen in die Falle, die so hohe Geschwindigkeit haben dass die schwache Abstoßung durch die Elektroden überhaupt keine Rolle spielt. Im Inneren der Falle werden die Ionen dann durch Laserstrahlung mit geeigneter Frequenz fast bis zum absoluten Nullpunkt abgekühlt (d. h. gebremst). Wenn man es schafft, sie fast zur Ruhe zu bringen, dann können sie nach dem Abschalten des Lasers die Falle nicht mehr verlassen, weil ihre thermische Energie dann nicht mehr ausreicht, um die elektrostatische Abstoßung durch die Elektroden zu überwinden.

Weil die beiden Ionen positiv geladen sind, stoßen sie sich auch wechselseitig ab. Ihr durchschnittlicher Abstand beträgt etwa $3 \mu\text{m}$, ist also etwa hundert mal kleiner als der Abstand zwischen den Elektroden, aber etwa 10 000 mal größer als der typische Abstand zwischen den beiden Atomen eines zweiatomigen Moleküls. Trotz dieses gewaltigen Abstands vibrieren die beiden Ionen relativ zueinander wie ein zweiatomiges Molekül, allerdings mit weitaus niedrigerer Frequenz. Dieses künstliche Molekül werden wir $\text{Be}^+ \& \text{Be}^+$ nennen.

Sein Termschema wird in Abb. 7.7 gezeigt. Der Energieabstand zwischen den Zuständen $|\downarrow\rangle_1 |\downarrow\rangle_2$ und $|\uparrow\rangle_1 |\uparrow\rangle_2$ beträgt etwa $h \cdot 2,5 \text{ GHz}$. Die Energie der beiden Zustände $|\uparrow\rangle_1 |\downarrow\rangle_2$ und $|\downarrow\rangle_1 |\uparrow\rangle_2$ liegt genau in der Mitte zwischen den beiden anderen Zuständen.

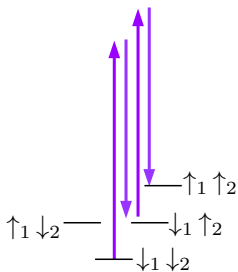


Abb. 7.7: Das Term-Schema von $\text{Be}^+ \& \text{Be}^+$. Die Energiedifferenz zwischen den Zuständen $|\downarrow\rangle_1 |\downarrow\rangle_2$ und $|\uparrow\rangle_1 |\uparrow\rangle_2$ beträgt etwa $h \cdot 2,5 \text{ GHz}$. Die violetten Pfeile sind nicht maßstabsgerecht gezeichnet!

Anfangs wird das künstliche Molekül Be⁺&Be⁺ durch Laserkühlung im Grundzustand $|\downarrow\rangle_1|\downarrow\rangle_2$ präpariert. Aus diesem wird es mithilfe von vier gekreuzten Laserstrahlen, deren Wellenlänge etwa 313 nm beträgt, durch doppelte stimulierte Raman-Streuung angeregt in den Zustand $|\uparrow\rangle_1|\uparrow\rangle_2$. Das klingt nicht nur ziemlich kompliziert, sondern ist es auch. Als Raman-Streuung wird ein Prozess bezeichnet, bei dem ein Atom oder Molekül ein Photon weder elastisch (ohne Energieänderung des Photons) streut noch komplett absorbiert, sondern das Photon zwar absorbiert aber gleichzeitig ein Photon mit nur geringfügig niedrigerer Energie emittiert. Dieser relativ unwahrscheinliche Vorgang kann dadurch „stimuliert“ (weniger unwahrscheinlich gemacht) werden, dass das Atom oder Molekül mit Photonen bestrahlt wird, die die gleiche Frequenz wie das Photon haben das emittiert wird.

In diesem Fall müssen sogar zwei Photonen inelastisch gestreut werden, was den Vorgang nochmals unwahrscheinlicher macht. Um die erforderliche Frequenzdifferenz von insgesamt 2,5 GHz zwischen den Laserstrahlen zustande zu bringen, muss deshalb mindestens einer der Laserstrahlen mit einem elektrooptischen Modulator präzise verstimmt werden.

Schließlich ist noch ein letzter Trick erforderlich: Das künstliche Molekül soll ja nicht im Zustand $|\uparrow\rangle_1|\uparrow\rangle_2$ präpariert werden, sondern im verschränkten Zustand

$$|\text{Be}^+\&\text{Be}^+\rangle \stackrel{(7.12)}{=} \sqrt{\frac{1}{2}} \left(|\downarrow\rangle_1|\downarrow\rangle_2 - |\uparrow\rangle_1|\uparrow\rangle_2 \right). \quad (7.14)$$

Man darf also nicht über das Ziel hinaus schießen, sondern muss die Anregung auf halbem Weg abbrechen.⁶¹ Wann der halbe Weg zurückgelegt ist, das hängt von der Stärke der Laser ab. Rowe et al. fanden heraus, dass sie die Laser nach ziemlich genau 0,5 μs

⁶¹ für Physiker: In der Theorie der Rabi-Oszillationen spricht man von einem $\pi/2$ -Puls. Eine elementare Einführung in diese Thematik findet man in [48].

abschalten mussten, um den verschränkten Zustand möglichst perfekt zu präparieren.

Zur Analyse des $\text{Be}^+ \& \text{Be}^+$ -Moleküls hätten wir gerne einen Detektor mit den vier Eigenvektoren

$$\begin{aligned} |\uparrow\rangle_{D1} |\uparrow\rangle_{D2} & & |\uparrow\rangle_{D1} |\downarrow\rangle_{D2} \\ |\downarrow\rangle_{D1} |\uparrow\rangle_{D2} & & |\downarrow\rangle_{D1} |\downarrow\rangle_{D2} . \end{aligned} \quad (7.15)$$

Wie die Stern-Gerlach-Magneten im Gedankenexperiment von Bohm und die polarisierenden Strahlteiler im Experiment von Aspect et al. sollen die Detektoren drehbar sein. Wenn der Detektor D1 um den Winkel φ_1 und der Detektor D2 um den Winkel φ_2 gedreht wird, dann soll beispielsweise aus dem Eigenvektor

$$|\uparrow\rangle_{D1} |\uparrow\rangle_{D2}$$

der Eigenvektor

$$(a |\uparrow\rangle_{D1} + b |\downarrow\rangle_{D1}) (c |\uparrow\rangle_{D2} + d |\downarrow\rangle_{D2})$$

werden, in dem die Amplituden a und b von φ_1 und die Amplituden c und d von φ_2 abhängen. Die Wahrscheinlichkeiten der verschiedenen Messergebnisse können dann wie gewohnt als Betragsquadrate der verschiedenen Projektionsamplituden berechnet werden. In diesen Berechnungen werden beispielsweise Terme der Art

$$\begin{aligned} & \left| (b_{D1} \langle \downarrow | + a_{D1} \langle \uparrow |) | \downarrow \rangle_1 \right|^2 = \\ & = \left| b \underbrace{\langle \downarrow | \downarrow \rangle_1}_1 + a \underbrace{\langle \uparrow | \downarrow \rangle_1}_0 \right|^2 = |b|^2 \end{aligned} \quad (7.16)$$

auftauchen. So einen drehbaren Detektor hatten Rowe et al. nicht. Aber sie überlegten sich, dass man genau die gleichen Betragsquadrate erhält, wenn man statt der Detektoren den Zustandsvektor

$|\text{Be}^+ \& \text{Be}^+\rangle$ dreht. Dann erhält man beispielsweise anstelle von (7.16) das Betragsquadrat

$$\begin{aligned} & \left| \text{D1} \langle \downarrow | (a | \uparrow \rangle_1 + b | \downarrow \rangle_1) \right|^2 = \\ & = \left| a \underbrace{\text{D1} \langle \downarrow | | \uparrow \rangle_1}_0 + b \underbrace{\text{D1} \langle \downarrow | | \downarrow \rangle_1}_1 \right|^2 = |b|^2. \end{aligned} \quad (7.17)$$

Es ist also egal, ob man den Detektor oder den Zustandsvektor des untersuchten Systems dreht, das Ergebnis ist in beiden Fällen das gleiche.

In Anhang A.5 wird erklärt, auf welche Weise Rowe et al. die Drehungen

$$| \uparrow \rangle_1 \xrightarrow{\text{Drehung um } \varphi_1} a_1 | \uparrow \rangle_1 + b_1 | \downarrow \rangle_1 \quad (7.18a)$$

$$| \downarrow \rangle_1 \xrightarrow{\text{Drehung um } \varphi_1} c_1 | \uparrow \rangle_1 + d_1 | \downarrow \rangle_1 \quad (7.18b)$$

$$| \uparrow \rangle_2 \xrightarrow{\text{Drehung um } \varphi_2} a_2 | \uparrow \rangle_2 + b_2 | \downarrow \rangle_2 \quad (7.18c)$$

$$| \downarrow \rangle_2 \xrightarrow{\text{Drehung um } \varphi_2} c_2 | \uparrow \rangle_1 + d_2 | \downarrow \rangle_2, \quad (7.18d)$$

der Zustände zustande brachten, in denen die Amplituden a_1, b_1, c_1, d_1 von einem variabel wählbaren Winkel φ_1 abhängen, und die Amplituden a_2, b_2, c_2, d_2 von einem variabel wählbaren Winkel φ_2 abhängen.

Durch Drehung um die Winkel φ_1 und φ_2 wird der Zustandsvektor (7.14) des $\text{Be}^+ \& \text{Be}^+$ -Moleküls zu

$$\begin{aligned} & |\text{Be}^+ \& \text{Be}^+\rangle \stackrel{(7.14), (7.18)}{=} \\ & = \sqrt{\frac{1}{2}} \left[(c_1 | \uparrow \rangle_1 + d_1 | \downarrow \rangle_1) (c_2 | \uparrow \rangle_1 + d_2 | \downarrow \rangle_2) - \right. \\ & \quad \left. - (a_1 | \uparrow \rangle_1 + b_1 | \downarrow \rangle_1) (a_2 | \uparrow \rangle_2 + b_2 | \downarrow \rangle_2) \right]. \end{aligned} \quad (7.19)$$

Dieser Zustandsvektor wird dann wie gewohnt auf die vier Eigenvektoren (7.15) der Messapparatur projiziert:

$$\begin{aligned}
 |\text{Be}^+ \& \text{Be}^+\rangle &= |7.19\rangle = \\
 &= |\uparrow\rangle_{D1} |\uparrow\rangle_{D2} D2 \langle \uparrow |_{D1} \langle \uparrow || 7.19 \rangle + \\
 &+ |\uparrow\rangle_{D1} |\downarrow\rangle_{D2} D2 \langle \downarrow |_{D1} \langle \uparrow || 7.19 \rangle + \\
 &+ |\downarrow\rangle_{D1} |\uparrow\rangle_{D2} D2 \langle \uparrow |_{D1} \langle \downarrow || 7.19 \rangle + \\
 &+ |\downarrow\rangle_{D1} |\downarrow\rangle_{D2} D2 \langle \downarrow |_{D1} \langle \downarrow || 7.19 \rangle \quad (7.20)
 \end{aligned}$$

Die jeweilige Wahrscheinlichkeit für jedes der vier möglichen Messergebnisse berechnet man nach der Born'schen Regel (5.19b) als Betragsquadrat der entsprechenden Projektionsamplitude:

$$W(\uparrow_1 \uparrow_2) = \left| D2 \langle \uparrow |_{D1} \langle \uparrow || 7.19 \rangle \right|^2 \quad (7.21a)$$

$$W(\uparrow_1 \downarrow_2) = \left| D2 \langle \downarrow |_{D1} \langle \uparrow || 7.19 \rangle \right|^2 \quad (7.21b)$$

$$W(\downarrow_1 \uparrow_2) = \left| D2 \langle \uparrow |_{D1} \langle \downarrow || 7.19 \rangle \right|^2 \quad (7.21c)$$

$$W(\downarrow_1 \downarrow_2) = \left| D2 \langle \downarrow |_{D1} \langle \downarrow || 7.19 \rangle \right|^2 \quad (7.21d)$$

Um diese Wahrscheinlichkeits-Vorhersagen der Quantentheorie experimentell zu prüfen, richteten Rowe et al. ihr Experiment vollautomatisch computergesteuert ein. Mit vier verschiedenen Kombinationen der Winkel φ_1 und φ_2 führten sie jeweils 20 000 Läufe des Experiments durch. Jeder der Läufe bestand aus drei Schritten:

- * Zuerst wurde das $\text{Be}^+ \& \text{Be}^+$ -Molekül im verschränkten Zustand (7.14) präpariert.
- * Dann wurden die gewünschten Phasenverschiebungen durchgeführt und damit das $\text{Be}^+ \& \text{Be}^+$ -Molekül in den Zustand (7.19) gedreht.

- * Schließlich wurde das $\text{Be}^+ \& \text{Be}^+$ -Molekül mit einem 1 ms langen „Detektionspuls“ bestrahlt, dessen Frequenz⁶² so gewählt wurde, dass eine intensive Resonanz-Lumineszenz auftrat wenn das bestrahlte Ion sich im Zustand $|\downarrow\rangle$ befand, aber keine Lumineszenz auftrat, wenn das bestrahlte Ion sich im Zustand $|\uparrow\rangle$ befand.

Die Ergebnisse von je 20 000 Durchläufen des Experiments mit zwei verschiedenen Winkel-Einstellungen sind in den Diagrammen 7.8 auf der nächsten Seite dargestellt. Auf der vertikalen Achse ist eingetragen, wie viele Lumineszenz-Photonen in einem Durchlauf des Experiments gezählt wurden, auf der waagerechten Achse ist eingetragen, in wie vielen Durchläufen die jeweilige Zahl von Photonen auftrat.

Man erkennt in den Diagrammen drei deutlich voneinander abgesetzte Gruppen. Die Experimentatoren interpretierten diese Ergebnisse folgendermaßen:

- * Wenn mehr als 85 Lumineszenzphotonen gezählt wurden, dann befanden sich offenbar beide Ionen im Zustand $|\downarrow\rangle$ (beide Ionen leuchten).
- * Wenn 25 bis 85 Lumineszenzphotonen gezählt wurden, dann befand sich offenbar eines der Ionen im Zustand $|\downarrow\rangle$ (ein Ion leuchtet), und ein Ion im Zustand $|\uparrow\rangle$ (ein Ion ist dunkel).
- * Wenn weniger als 25 Lumineszenzphotonen gezählt wurden, dann befanden sich offenbar beide Ionen im Zustand $|\uparrow\rangle$ (beide Ionen sind dunkel).

⁶² Die Frequenz war exakt auf den Übergang zwischen dem Grundzustand \downarrow und dem Zustand ${}^2\text{P}_{3/2}$ der Einzel-Ionen einjustiert, siehe das Termschema A.4. Ein Ion, das sich beim Beginn des Detektionspulses im Zustand $|\downarrow\rangle$ befindet, wird bei dieser Bestrahlung zwischen den Zuständen $|\downarrow\rangle$ und $|\text{}^2\text{P}_{3/2}\rangle$ oszillieren, wobei eine intensive Resonanz-Lumineszenz auftritt, die leicht beobachtet werden kann.

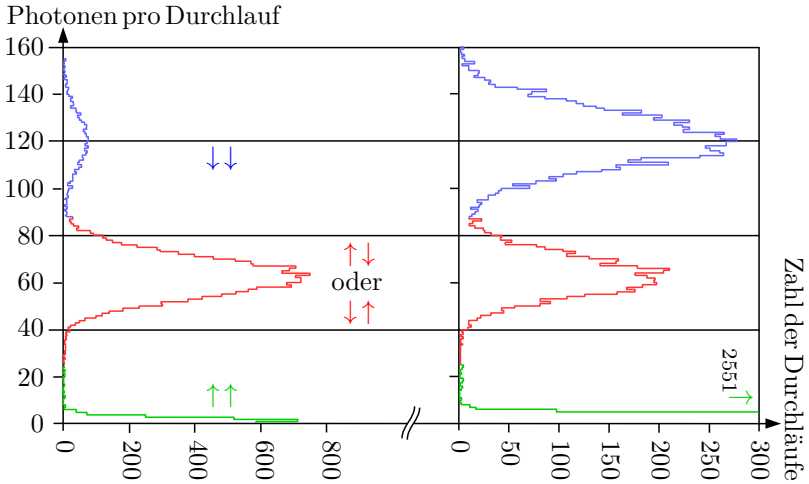


Fig. 7.8: Linkes Diagramm: Das Ergebnis von 20 000 Durchläufen des Experiments mit $\varphi_1 = +67,5^\circ \hat{=} +3\lambda/16$, $\varphi_2 = +67,5^\circ \hat{=} +3\lambda/16$
 Rechtes Diagramm: Das Ergebnis von 20 000 Durchläufen des Experiments mit $\varphi_1 = +67,5^\circ \hat{=} +3\lambda/16$, $\varphi_2' = -22,5^\circ \hat{=} -\lambda/16$

Wir definieren

$N_{\downarrow\downarrow}$ als die Anzahl der Durchläufe, bei denen mehr als 85 Photonen gezählt wurden.

$N_{\uparrow\downarrow}$ oder $\downarrow\uparrow$ als die Anzahl der Durchläufe, bei denen 25 bis 85 Photonen gezählt wurden.

$N_{\uparrow\uparrow}$ als die Anzahl der Durchläufe, bei denen weniger als 25 Photonen gezählt wurden.

Aus diesen Werten ergeben sich die Wahrscheinlichkeiten (d. h. die relativen Häufigkeiten), mit denen die verschiedenen Zustände im Experiment auftraten:

$$W(\uparrow_1 \uparrow_2) = \frac{N_{\uparrow\uparrow}}{N_{\uparrow\uparrow} + N_{\uparrow\downarrow \text{ oder } \downarrow\uparrow} + N_{\downarrow\downarrow}} \quad (7.22a)$$

$$W(\uparrow_1 \downarrow_2) + W(\downarrow_1 \uparrow_2) = \frac{N_{\uparrow\downarrow \text{ oder } \downarrow\uparrow}}{N_{\uparrow\uparrow} + N_{\uparrow\downarrow \text{ oder } \downarrow\uparrow} + N_{\downarrow\downarrow}} \quad (7.22b)$$

$$W(\downarrow_1 \downarrow_2) = \frac{N_{\downarrow\downarrow}}{N_{\uparrow\uparrow} + N_{\uparrow\downarrow \text{ oder } \downarrow\uparrow} + N_{\downarrow\downarrow}} . \quad (7.22c)$$

Rowe et al. bestimmten diese drei Wahrscheinlichkeiten für die vier Winkel-Kombinationen

$$\begin{aligned} \varphi_1 &= +67,5^\circ \hat{=} + \frac{3}{16} \lambda \\ \varphi_2 &= +67,5^\circ \hat{=} + \frac{3}{16} \lambda \\ \varphi'_1 &= -22,5^\circ \hat{=} - \frac{1}{16} \lambda \\ \varphi'_2 &= -22,5^\circ \hat{=} - \frac{1}{16} \lambda . \end{aligned} \quad (7.23)$$

Aus diesem Ergebnis ergeben sich die Korrelationsfunktionen

$$\overline{r_{\varphi_1} \cdot r_{\varphi_2}} = W(\uparrow_1 \uparrow_2) + W(\downarrow_1 \downarrow_2) - W(\uparrow_1 \downarrow_2) - W(\downarrow_1 \uparrow_2) ,$$

mit denen man die Bell'sche Ungleichung (6.16)

$$-2 \leq \bar{q} \leq +2 \quad (7.24)$$

$$\text{mit } \bar{q} = \overline{r_{\varphi_1} \cdot r_{\varphi_2}} + \overline{r_{\varphi_1} \cdot r_{\varphi'_2}} + \overline{r_{\varphi'_1} \cdot r_{\varphi_2}} - \overline{r_{\varphi'_1} \cdot r_{\varphi'_2}}$$

prüfen kann. Zur Erinnerung: Die Ungleichung (7.24) muss bei beliebigen Einstellungen $\varphi_1, \varphi'_1, \varphi_2, \varphi'_2$ erfüllt sein, falls die Annahme **B** zutrifft. Die Annahme **B** besagt im Fall des künstlichen Be⁺&Be⁺-Moleküls, dass die Ionen bereits vor der Messung in einem der Zustände $|\uparrow\rangle_1$ oder $|\downarrow\rangle_1$ bzw. $|\uparrow\rangle_2$ oder $|\downarrow\rangle_2$ sind, und dass dieser Zustand durch die Messung lediglich festgestellt wird. Dagegen nimmt die Quantentheorie an, dass die Zustände $|\uparrow\rangle_1$

oder $|\downarrow\rangle_1$ bzw. $|\uparrow\rangle_2$ oder $|\downarrow\rangle_2$ erst durch die Messung erschaffen werden. Solange das $\text{Be}^+ \& \text{Be}^+$ -Molekül im verschränkten Zustand (7.14) existiert, haben die einzelnen Ionen laut Quantentheorie keine eigenständige Existenz und deshalb auch keinen Zustand.

Wenn man \bar{q} mithilfe der Quantentheorie berechnet, dann wird die Bell'sche Ungleichung (7.24) bei vielen Winkeln verletzt, und zwar besonders deutlich bei den von Rowe et al. untersuchten Winkeln (7.23). Bei diesen Winkeln berechnet man mit der Quantentheorie

$$\bar{q}_{\text{Quantentheorie}} = 2\sqrt{2} \approx 2,83 . \quad (7.25)$$

Rowe et al. fanden

$$\bar{q}_{\text{Experiment}} = 2,25 \pm 0,03 + 0,12 . \quad (7.26)$$

Das liegt deutlich über 2, und widerlegt damit Annahme B. Der experimentelle Wert liegt aber auch deutlich unter dem Wert $2\sqrt{2}$ der Quantentheorie. Warum? Dafür sahen Rowe et al. hauptsächlich drei Gründe:

- * Die Aufnahme eines kompletten Datensatzes – insgesamt 80 000 Durchläufe des Experiments mit vier verschiedenen Kombinationen der Winkel (7.23) – dauerte etwa vier Minuten. Das klingt nach wenig, es erwies sich aber als sehr schwierig die Geräte, mit denen die Phasenverschiebungen gesteuert wurden, über den gesamten Zeitraum stabil zu halten. Die Ungenauigkeiten bei der Einstellung der Phasenverschiebungen führten zum Fehler $\pm 0,03$ in (7.26).
- * Bei der Detektion wurden ungewollt einige Ionen aus dem Zustand $|\downarrow\rangle$ durch den Detektions-Puls in den Zustand $|\uparrow\rangle$ angeregt, waren dann dunkel, und wurden deshalb falsch bewertet. Rowe et al. schätzten, dass ihr Ergebnis um den in (7.26) angegebenen Wert von $+0,12$ höher wäre, wenn dieser

Fehler bei der Detektion abgestellt werden könnte.

- * Ein dritter Fehler bestand darin, dass die Präparation des künstlichen Moleküls $\text{Be}^+\&\text{Be}^+$ im verschränkten Zustand (7.14) tatsächlich nur in 88 % der Fälle gelang. Die Experimentatoren konnten die Auswirkungen dieses Fehlers auf (7.26) nicht quantifizieren. Es ist aber nicht zu bezweifeln, dass eine Verbesserung der Präparation den Wert (7.26) noch weiter erhöhen würde.

Trotz all dieser Schwierigkeiten wird die Annahme **B** durch das Experiment von Rowe et al. zweifelsfrei widerlegt.

Dieses Experiment ist zum einen deshalb bemerkenswert, weil es – anders als alle anderen in diesem Buch beschriebenen Experimente – nicht auf der Verschränkung der Polarisation von Photonenpaaren beruht, sondern auf der Verschränkung der Spins von Atomen (der aufs Engste mit dem magnetischen Moment zusammenhängt), und damit das Gedankenexperiment von Bohm realisiert (auch wenn die Verschränkung in diesem Fall nicht mit Stern-Gerlach-Magneten sondern mit spektroskopischen Methoden beobachtet wurde).

Das Experiment ist zum anderen auch deshalb bemerkenswert weil es nicht auf die Annahme des “fair sampling” angewiesen ist. Damit ist folgendes gemeint: Photonendetektoren haben in der Regel eine Effizienz von nur etwa 10%, d. h. sie sprechen nur bei etwa einem von zehn Photonen überhaupt an. Bei den Experimenten mit zwei verschränkten Photonen kann die Korrelation nur überprüft werden wenn *beide* Photonen detektiert werden, und das geschieht nur bei etwa einem von hundert Photonenpaaren. Man muss also darauf vertrauen, dass das eine Prozent beobachteter Photonenpaare auch wirklich repräsentativ für die Gesamtheit aller Photonenpaare ist. Die Annahme des “fair sampling” ist zwar sehr plausibel, und wird von niemandem ernsthaft angezweifelt. Trotzdem hätte man natürlich gerne Gewissheit.

Das Experiment von Rowe et al. schließt diese Lücke in der Be-

weiskette. Denn in diesem Experiment werden ja bei jedem einzelnen Durchlauf Dutzende von Photonen registriert, siehe Abb. 7.8, so dass es überhaupt nichts ausmacht wenn ein großer Teil der Photonen unbemerkt bleiben sollte. Dieser Umstand war den Autoren so wichtig, dass sie im Titel “Experimental violation of a Bell’s inequality with efficient detection” ihrer Veröffentlichung [47] ausdrücklich darauf hinwiesen.

7.3 Verschränkte SPDC-Photonen

In Abschnitt 3.5 haben wir das SPDC-Verfahren (SPDC = spontaneous parametric downconversion zur Herstellung von Photonenpaaren kennen gelernt, siehe Abb. 3.6 auf Seite 62).

Die Trajektorien der beiden Tochterphotonen und des Pumpphotons liegen bei diesem Verfahren stets in einer Ebene, weil nur so der Impuls erhalten werden kann. Diese Ebene braucht aber für verschiedene Paare von Tochterphotonen nicht die gleiche zu sein. Bei SPDC Typ I liegen die Trajektorien der Tochterphotonen auf einem Kegelmantel, dessen Achse durch den Pumpstrahl definiert wird, siehe Abb. 3.6 auf Seite 62. Bei SPDC Typ II liegen die Trajektorien der Tochterphotonen dagegen auf unterschiedlichen Kegelmänteln, wie in Abb. 7.9 skizziert. Ob SPDC Typ I oder SPDC Typ II auftritt hängt davon ab, wie der Kristall relativ zum Pumplicht-Strahl justiert wird.

Wenn das Pumpphoton L_{90} -polarisiert ist, wird beim SPDC-II-Prozess das Tochterphoton auf dem einen Kegelmantel im Zustand $|L_0\rangle$ präpariert, das Tochterphoton auf dem anderen Kegelmantel im Zustand $|L_{90}\rangle$.

Wenn man – wie in Abb. 7.10 skizziert – mithilfe von Lochblenden Photonen auswählt, deren Trajektorien genau auf den Schnittlinien der beiden Kegelmäntel liegen, dann kann man das Gesamtsystem der beiden Tochterphotonen im Zustand

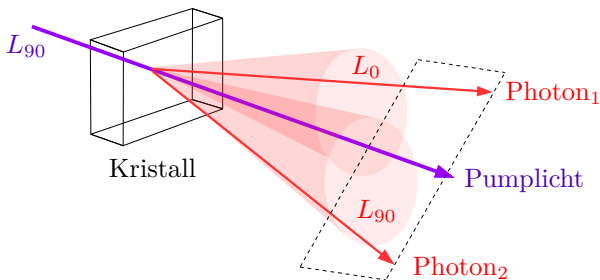


Fig. 7.9: Bei SPDC Typ II liegen die Trajektorien der beiden Tochterphotonen (rote Pfeile) auf unterschiedlichen Kegelmänteln, aber stets symmetrisch zum Pumpstrahl, wie durch die gestrichelte Ebene angedeutet.

$$|\text{Photonenpaar}\rangle = \sqrt{\frac{1}{2}} \left(|L_0\rangle_1 |L_{90}\rangle_2 - |L_{90}\rangle_1 |L_0\rangle_2 \right) \quad (7.27)$$

präparieren.⁶³ Wenn dies der Zustandsvektor des Photonenspaars ist, dann hat – solange die Polarisation noch nicht gemessen wurde – laut Quantentheorie weder Photon₁ noch Photon₂ irgendeine Polarisation. Dennoch weiß man mit Sicherheit: Eine künftige Polarisationsmessung wird die beiden Photonen in Polarisationszuständen präparieren, bei denen die in (7.27) beschriebene Korrelation zwischen den Polarisationen von Photon₁ und Photon₂ realisiert wird.

Das ist etwas wesentlich anderes als die von EPR favorisierte Annahme **B**, dass die Photonen eine Polarisation haben, die uns lediglich unbekannt ist. Indem man mithilfe der Photonenspaare (7.27) die Bell'sche Ungleichung prüft, kann man die Annahme **B** experimentell widerlegen.

Das taten im Jahr 1995 Kwiat, Mattle, Weinfurter, Zeilinger, Ser-

⁶³ Durch geringfügige Variation der Justierung kann man auch andere verschränkte Zustände präparieren, siehe [49]. Wir werden uns aber in diesem Kapitel auf die Untersuchung des Zustands (7.27) konzentrieren.

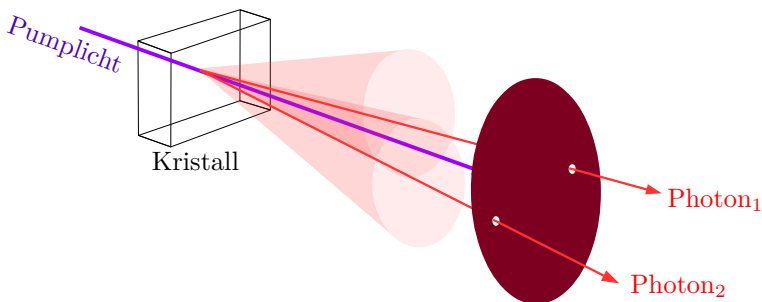


Fig. 7.10: Präparation des verschränkten Zustands (7.27)

gienko, und Shih [49]. Sie erzeugten Paare verschränkter Photonen durch SPDC Typ II mithilfe eines geeignet justierten Kristalls aus β -Bariumborat (BBO). Wie in Abb. 7.10 gezeichnet justierten sie die beiden Lochblenden, mit denen die Photonenpaare ausgewählt wurden, so, dass sowohl Photon_1 als auch Photon_2 auf *beiden* Kegelmänteln lagen. Dadurch wurden die erzeugten Photonenpaare im Zustand (7.27) präpariert.

Weitere Details ihres Experiments sind in Abb. 7.11 abgebildet. Sowohl Photon_1 als auch Photon_2 liefen durch je ein $\lambda/2$ -Plättchen (sprich: lambda-halbe-Plättchen). Durch Drehung des $\lambda/2$ -Plättchens kann man die Polarisationssebene der Photonen beliebig

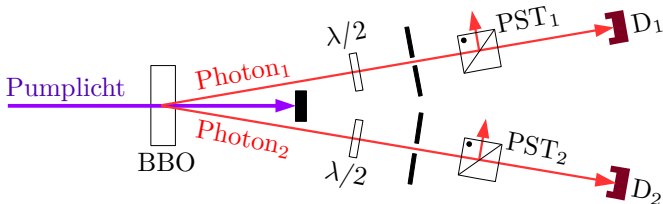


Fig. 7.11: Das Experiment von Kwiat et al. [49]

drehen.⁶⁴ Wenn man die Polarisationssebene des Photons mit dem $\lambda/2$ -Plättchen um den Winkel $+\gamma$ dreht, dann hat das auf die Zählrate des Detektors genau die gleiche Wirkung wie eine Drehung des polarisierenden Strahlteilers PST um den Winkel $-\gamma$. Die $\lambda/2$ -Plättchen wurden in diesem Experiment also nur eingesetzt, damit man die Strahlteiler nicht zu drehen brauchte.

Wenn die Polarisation von Photon_1 mithilfe eines $\lambda/2$ -Plättchens um γ_1 gedreht wird, und die Polarisation von Photon_2 mithilfe eines $\lambda/2$ -Plättchens um γ_2 gedreht wird, dann wird der Zustandsvektor des Photonenpaares folgendermaßen geändert:

$$\begin{aligned} |7.27\rangle &= \sqrt{\frac{1}{2}} \left(|L_0\rangle_1 |L_{90}\rangle_2 - |L_{90}\rangle_1 |L_0\rangle_2 \right) \xrightarrow{\lambda/2\text{-Plättchen}} \\ &\longrightarrow \sqrt{\frac{1}{2}} \left(|L_{0+\gamma_1}\rangle_1 |L_{90+\gamma_2}\rangle_2 - |L_{90+\gamma_1}\rangle_1 |L_{0+\gamma_2}\rangle_2 \right) \quad (7.28) \end{aligned}$$

In diesem Zustand traf das Photonenpaar auf die polarisierenden Strahlteiler PST_1 und PST_2 . Beide Strahlteiler waren so justiert, dass sie Licht mit L_0 -Polarisation (in der Papierebene von Zeichnung 7.11 polarisiertes Licht) transmittierten, und Licht reflektierten das L_{90} -polarisiert war (Polarisation senkrecht zur Papierebene von Zeichnung 7.11).

Reflektierte Photonen wurden ignoriert, transmittierte Photonen wurden mit den Detektoren D_1 und D_2 detektiert. Wenn nur ein Detektor ansprach, wurde das Ereignis ebenfalls ignoriert. Nur Koinzidenzen (beide Detektoren sprechen an) wurden gezählt. Weil beide Strahlteiler transmittierte Photonen im Zustand $|L_0\rangle$ präparieren, hat die Messapparatur nur den einen Eigenvektor

$$|L_0\rangle_1 |L_0\rangle_2 . \quad (7.29)$$

Die Projektion von (7.28) auf diesen Eigenvektoren der Messapparatur ergibt:

⁶⁴ In Anhang A.6 wird erklärt wie das $\lambda/2$ -Plättchen funktioniert.

$$\begin{aligned}
 |7.28\rangle &= \sqrt{\frac{1}{2}} \left[|L_0\rangle_1 |L_0\rangle_2 \cdot \right. \\
 &\quad \cdot \left. \left({}_1\langle L_0 || L_{\gamma_1} \rangle_1 {}_2\langle L_0 || L_{90+\gamma_2} \rangle_2 - {}_1\langle L_0 || L_{90+\gamma_1} \rangle_1 {}_2\langle L_0 || L_{\gamma_2} \rangle_2 \right) \right]
 \end{aligned}$$

Die Wahrscheinlichkeit $W_{TT}(\gamma_1, \gamma_2)$ dafür, dass beide Photonen transmittiert werden, ist nach der Born'schen Regel (5.19b) gleich dem Betragsquadrat der entsprechenden Projektionsamplitude:

$$\begin{aligned}
 W_{TT} &= \frac{1}{2} \left| \langle L_0 || L_{\gamma_1} \rangle \langle L_0 || L_{90+\gamma_2} \rangle - \langle L_0 || L_{90+\gamma_1} \rangle \langle L_0 || L_{\gamma_2} \rangle \right|^2 = \\
 &= \frac{1}{2} \left| \cos(\gamma_1) \cos(90^\circ + \gamma_2) - \cos(90^\circ + \gamma_1) \cos(\gamma_2) \right|^2 = \\
 &= \frac{1}{2} \sin^2(\gamma_1 - \gamma_2) \tag{7.30}
 \end{aligned}$$

Die letzte Umformung kann man der Formelsammlung entnehmen oder das Ergebnis einfach glauben.

Kwiat et al. stellten mithilfe der $\lambda/2$ -Plättchen die Polarisation von Photon₁ stets auf $\gamma_1 = -45^\circ$ ein, während die Polarisation von Photon₂ in verschiedenen Durchläufen des Experiments auf verschiedene Winkel γ_2 eingestellt wurde. In Abb. 7.12 auf der nächsten Seite ist $W_{TT} = (7.30)$ als Funktion von γ_2 (bei $\gamma_1 = -45^\circ$) als gelbe Kurve eingezeichnet.⁶⁵ Es gilt die rechte Skala, nicht die linke!

Bei 12 verschiedenen Einstellungen von γ_2 zählten Kwiat et al. jeweils 200 Sekunden lang, wie oft die Detektoren D₁ und D₂ (siehe Abb. 7.11) gleichzeitig⁶⁶ ansprachen. Die Anzahl dieser Koinzidenzen ist in Abb. 7.12 durch rote Punkte dargestellt. Es gilt die linke

⁶⁵ Im Artikel von Kwiat et al. [49, FIG.3] wurden offensichtlich die Kurven HV-VH und HV+VH irrtümlich vertauscht. Unsere Abb. 7.12 ist korrekt.

⁶⁶ Kwiat et al. haben vergessen in ihrer Veröffentlichung [49] zu dokumentieren, wie groß der Abstand der Photonen sein durfte, damit ihre Detektion als „gleichzeitig“ anerkannt wurde. Typischerweise wählen Experimentatoren dieses Zeitfenster 2 bis 10 Nanosekunden groß.

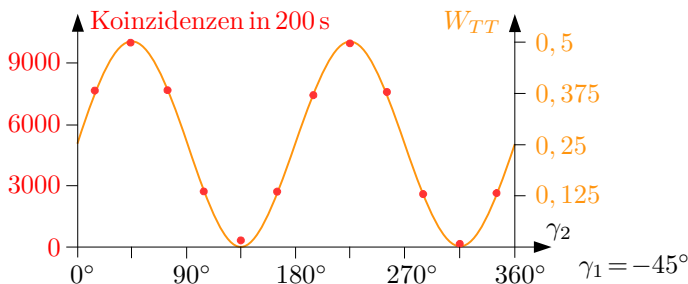


Fig. 7.12: Das Experiment von Kwiat et al.[49]

Skala, die so gestaucht wurde, dass die roten Punkte möglichst genau auf der gelben Linie liegen.

Es fällt auf, dass die Koinzidenzen bei 135° und bei 315° nicht exakt Null sind, wie es laut der gelben Linie – spricht laut Quantentheorie – eigentlich sein sollte. Experimente sind halt niemals perfekt. Trotzdem kann man sicherlich feststellen, dass die Quantentheorie sehr gut mit dem experimentellen Ergebnis übereinstimmt.

Die Annahme **B** wird durch dieses Ergebnis widerlegt, denn die Bell'sche Ungleichung wird verletzt. Zur Erinnerung: Die Bell'sche Ungleichung (6.16) lautet:

$$-2 \leq \bar{q} \leq +2 \quad (7.31)$$

$$\text{mit } \bar{q} = \overline{r_{\gamma_1} \cdot r_{\gamma_2}} + \overline{r_{\gamma_1} \cdot r_{\gamma'_2}} + \overline{r_{\gamma'_1} \cdot r_{\gamma_2}} - \overline{r_{\gamma'_1} \cdot r_{\gamma'_2}}$$

$$\text{mit } \overline{r_{\gamma_1} \cdot r_{\gamma_2}} = W_{TT} + W_{RR} - W_{TR} - W_{RT}$$

Wenn Annahme **B** richtig wäre, dann müsste diese Ungleichung bei beliebigen Winkeln $\gamma_1, \gamma_2, \gamma'_1, \gamma'_2$ gelten.

Mit der Wahrscheinlichkeit $W_{TT}(\gamma_1, \gamma_2)$ wird Photon₁ vom polarisierenden Strahlteiler PST₁ transmittiert, wenn es mit Polarisation L_{γ_1} einläuft, und gleichzeitig Photon₂ vom polarisierenden Strahlteiler PST₂ transmittiert, wenn es mit Polarisation L_{γ_2} einläuft,

mit $W_{TR}(\gamma_1, \gamma_2)$ wird Photon₁ transmittiert und Photon₂ reflektiert,

mit $W_{RT}(\gamma_1, \gamma_2)$ wird Photon₁ reflektiert und Photon₂ transmittiert,

mit $W_{RR}(\gamma_1, \gamma_2)$ wird Photon₁ reflektiert und Photon₂ reflektiert. Diese Wahrscheinlichkeiten hängen mit der Anzahl transmittierter und reflektierter Photonen folgendermaßen zusammen:

$$W_{TT} = \frac{N_{TT}}{N_{TT} + N_{RR} + N_{TR} + N_{RT}} \quad (7.32a)$$

$$W_{RR} = \frac{N_{RR}}{N_{TT} + N_{RR} + N_{TR} + N_{RT}} \quad (7.32b)$$

$$W_{TR} = \frac{N_{TR}}{N_{TT} + N_{RR} + N_{TR} + N_{RT}} \quad (7.32c)$$

$$W_{RT} = \frac{N_{RT}}{N_{TT} + N_{RR} + N_{TR} + N_{RT}} \quad (7.32d)$$

N_{TT} mal wurde Photon₁ transmittiert und Photon₂ transmittiert,
 N_{TR} mal wurde Photon₁ transmittiert und Photon₂ reflektiert,
 N_{RT} mal wurde Photon₁ reflektiert und Photon₂ transmittiert,
 N_{RR} mal wurde Photon₁ reflektiert und Photon₂ reflektiert.

Tatsächlich registrierten Kwiat et al. nur transmittierte Photonen, die reflektierten Photonen wurden nicht registriert. Sie konnten die anderen Werte aber aus

$$N_{RT}(\gamma_1, \gamma_2) = N_{TT}(\gamma_1 + 90^\circ, \gamma_2)$$

$$N_{TR}(\gamma_1, \gamma_2) = N_{TT}(\gamma_1, \gamma_2 + 90^\circ)$$

$$N_{RR}(\gamma_1, \gamma_2) = N_{TT}(\gamma_1 + 90^\circ, \gamma_2 + 90^\circ)$$

erschließen. Deswegen zählten sie die Koinzidenzen N_{TT} bei folgenden Winkeln

$$\gamma_1 = -22,5^\circ \quad \gamma_1^\perp = -22,5^\circ + 90^\circ = +67,5^\circ \quad (7.33a)$$

$$\gamma_1' = +22,5^\circ \quad \gamma_1'^\perp = +22,5^\circ + 90^\circ = +112,5^\circ \quad (7.33b)$$

$$\gamma_2 = -45^\circ \quad \gamma_2^\perp = -45^\circ + 90^\circ = +45^\circ \quad (7.33c)$$

$$\gamma_2' = 0^\circ \quad \gamma_2'^\perp = 0^\circ + 90^\circ = +90^\circ, \quad (7.33d)$$

und berechneten daraus den Wert

$$\bar{q} = -2,649 \pm 0,006, \quad (7.34)$$

der die Bell'sche Ungleichung (7.31) um

$$\frac{2,649 - 2}{0,006} \approx 108$$

Standardabweichungen⁶⁷ verletzt. Damit ist definitiv bewiesen, dass Annahme B falsch ist.

7.4 Raumartig getrennte Detektoren

Das Experiment von Rowe et al. das in Abschnitt 7.2 beschrieben wurde, hat das Schlupfloch der unzulänglichen Detektoreffizienz geschlossen, es ist auf die „fair sampling“ Hypothese nicht mehr angewiesen. Ein zweites Schlupfloch, das den Physikern noch mehr Sorgen machte, wird unter dem Namen „locality loophole“ oder „communication loophole“ diskutiert. Kann man sich wirklich sicher sein, dass nicht die beiden Stern-Gerlach-Magneten in Bohms Gedankenexperiment auf irgend eine unbekannte Weise herausfinden, auf welchen Winkel der jeweils andere Magnet eingestellt ist, und dann die Atome so ablenken, dass sich die passende Korrelation der beiden Messergebnisse ergibt? Oder dass in den Experimenten mit korrelierten Photonenpaaren die polarisierenden Strahlteiler und $\lambda/2$ -Plättchen irgendwie die Einstellung des jeweils

⁶⁷ Was eine „Standardabweichung“ ist, wurde bei Gleichung (3.14) erklärt.

anderen in Erfahrung bringen, und dann die Photonen so transmittieren oder reflektieren, dass die passende Korrelation beobachtet wird? Gewiss, niemand weiß auf welche Weise die Magnete und Strahlteiler, selbst wenn sie diese Informationen hätten, ihre Aktionen koordinieren könnten. Trotzdem ist dies ein Schlupfloch, das geschlossen werden sollte.

Bell hatte schon in seinem Artikel von 1964 erklärt, wie man das machen kann: Man muss die Einstellungen der Magnete bzw. Strahlteiler und $\lambda/2$ -Plättchen so kurz vor dem Eintreffen der Teilchen verändern, dass die Information über die aktuelle Einstellung – vorausgesetzt dass sie nicht schneller als mit Lichtgeschwindigkeit übertragen wird – unmöglich beim anderen Detektor eintreffen kann, bevor die Messung an beiden Detektoren abgeschlossen ist. In der Sprache der Relativitätstheorie sagt man, dass die Detektoren „raumartig voneinander getrennt“ sein müssen.

Ein Experiment mit raumartig getrennten Detektoren wurde erstmals 1998 von Weihs et al. realisiert [50]. Der Aufbau des Experiments ist in Abb. 7.13 auf der nächsten Seite skizziert. Es handelt sich um eine Weiterentwicklung des Experiments Abb. 7.11 von Kwiat et al. (beide Experimente wurden im gleichen Labor an der Universität Innsbruck durchgeführt). Die Photonenpaare wurden per SPDC Typ II mit einem BBO-Kristall (β -Bariumborat) erzeugt, und – wie in Abb. 7.10 skizziert – so ausgewählt, dass sie im verschränkten Zustand

$$|\text{Photonenpaar}\rangle = \sqrt{\frac{1}{2}} \left(|L_0\rangle_1 |L_{90}\rangle_2 - |L_{90}\rangle_1 |L_0\rangle_2 \right) \quad (7.35)$$

präpariert wurden. Weihs et al. detektierten nicht nur die transmittierten Photonen, sondern auch die reflektierten Photonen, maßen also – anders als Kwiat et al. – nicht nur N_{TT} , sondern direkt auch N_{TR} , N_{RT} , und N_{RR} .

Besonders wichtig waren zwei Veränderungen: Erstens wurden die $\lambda/2$ -Plättchen durch elektro-optische Modulatoren (EOM) ersetzt.

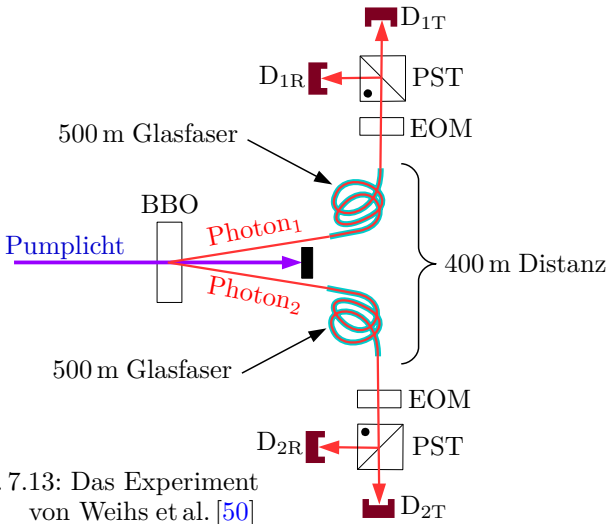


Abb. 7.13: Das Experiment von Weihs et al. [50]

Das sind Kristalle, die die Polarisation von Licht um einen mehr oder weniger großen Winkel drehen, wenn eine mehr oder weniger große elektrische Spannung an den Kristall gelegt wird. Welche Spannung angelegt wurde, d. h. um welche Winkel die Polarisationen von Photon₁ und Photon₂ gedreht wurden, das wurde von Zufallsgeneratoren gesteuert. Zweitens waren die Zufallsgeneratoren, elektrooptischen Modulatoren (EOM), polarisierenden Strahlteiler (PST) und Detektoren, die Photon₁ verarbeiteten, 400 m entfernt von den entsprechenden Geräten, die sich mit Photon₂ beschäftigten. Die Quelle der Photonenpaare befand sich etwa in der Mitte. Photon₁ und Photon₂ wurden durch je eine 500 m lange, teilweise aufgewickelte Glasfaser vom BBO-Kristall zur jeweiligen Detektorstation geführt.

Um eine Information mit Lichtgeschwindigkeit vom einen zum anderen Detektor zu übertragen, braucht man

$$\frac{400 \text{ m}}{c} = \frac{400 \text{ m}}{3 \cdot 10^8 \text{ m/s}} = 1,33 \mu\text{s} . \quad (7.36)$$

Das Zeitintervall, das die Experimentatoren benötigten, um mit schnellen physikalischen Zufalls-Generatoren Winkel für die Polarisatoren zu wählen, mithilfe der elektrooptischen Modulatoren diese Einstellungen zu realisieren, und schließlich das Messergebnis zu registrieren, lag deutlich unter den maximal zulässigen 1,3 Mikrosekunden. Das „locality“ Schlupfloch war also bei diesem Experiment zuverlässig geschlossen.

Jedes mal wenn einer der vier Detektoren ansprach, wurde dokumentiert zu welcher Zeit dies geschah (das wurde mit synchronisierten Atomuhren gemessen) und auf welche Polarisation der elektrooptische Modulator dabei eingestellt war. Nach Abschluss der Messungen wurden die Ereignisse dann einander zugeordnet: Wenn der Zeitabstand kleiner als 6 ns war wurde angenommen, dass es sich um Photon_1 und Photon_2 eines korrelierten Paares (7.35) handelte.

Weil die Zufallsgeneratoren in diesem Experiment je etwa 40 verschiedene Winkeleinstellungen der elektrooptische Modulatoren auswählten statt sich auf wenige Winkel zu konzentrieren, eigneten sich die Daten nicht für eine Auswertung der Bell'schen Ungleichung. Die Experimentatoren beschränkten sich deshalb darauf, aus ihren Daten Messkurven nach Art der roten Punkte in Abb. 7.12 auf Seite 188 herzustellen, und die gute Übereinstimmung dieses Resultats mit der Vorhersage der Quantentheorie festzustellen.

Damit war das locality Schlupfloch geschlossen. Leider war aber in diesem Experiment wegen der schlechten Effizienz der Photonen-Detektoren wieder die fair sampling Hypothese erforderlich.

Experimente, bei denen alle bekannten Schlupflöcher gleichzeitig geschlossen wurden, gelangen erst in den Jahren 2015 und 2016, dann aber gleich in vier verschiedenen Laboratorien [51–54]. Ich werde diese Experimente hier nicht im Einzelnen diskutieren und gebe nur bekannt, dass sie – wie von niemandem anders erwartet – die Bell'sche Ungleichung verletzen und damit die Alternative B erneut widerlegten.

7.5 Schlussfolgerungen

Was bedeuten die experimentellen Ergebnisse, die in den vorangegangenen Abschnitten beschrieben wurden? Fassen wir unsere Überlegungen nochmal übersichtlich zusammen:

- * Laut Quantentheorie können zwei Teilsysteme 1 und 2 zu einem Gesamtsystem 1&2 mit dem Zustandsvektor

$$|1&2\rangle = u|A\rangle_1|B\rangle_2 + v|B\rangle_1|A\rangle_2 \quad \text{mit } |u|^2 + |v|^2 = 1$$

verschränkt sein. Charakteristisch für verschränkte Zustandsvektoren ist, dass sie nicht als Produkt $|1\rangle|2\rangle$ von Zustandsvektoren

$$|1\rangle = a|A\rangle_1 + b|B\rangle_1 \quad \text{mit } |a|^2 + |b|^2 = 1$$

$$|2\rangle = c|A\rangle_2 + d|B\rangle_2 \quad \text{mit } |c|^2 + |d|^2 = 1$$

der Teilsysteme geschrieben werden können.

- * Wenn der Zustandsvektor des Gesamtsystems verschränkt ist, dann ordnet die Quantentheorie also nur dem Gesamtsystem einen Zustandsvektor zu, aber nicht den Teilsystemen. Das bedeutet: Die Teilsysteme existieren (laut Quantentheorie) nur als Bestandteile des Gesamtsystems, haben aber keine eigenständige Existenz.
- * Insbesondere kann das verschränkte Gesamtsystem Eigenschaften haben (z. B. eine bestimmte Polarisierung, oder ein bestimmtes magnetisches Moment in Richtung einer bestimmten Raumachse, oder auch z. B. einen bestimmten Ort oder einen bestimmten Impuls), die die Teilsysteme nicht haben.
- * Einstein, Podolski, und Rosen (EPR) wollten beweisen, dass die Quantentheorie in dieser Hinsicht unvollständig ist, dass nämlich auch die Teilsysteme alle diese Eigenschaften bereits

vor der Messung haben, dass diese Eigenschaften der Teilsysteme bei einer Messung also nicht erschaffen werden, sondern lediglich festgestellt werden.

- * Der vermeintliche Beweis von EPR erwies sich bei genauerer Analyse als Rohrkrepiierer. Die Verletzungen der Bell'schen Ungleichung sind⁵⁶ experimentelle Beweise dafür, dass die Korrelationen zwischen den Teilsystemen verschränkter Gesamtsysteme stärker sind als sie es maximal sein könnten, wenn die von EPR favorisierte Annahme **B** richtig wäre.

Überlegen wir nochmal ganz genau, was es mit der Bell'schen Ungleichung und ihrer Verletzung in den Experimenten auf sich hat. Wir haben aus den drei Annahmen

A1_{Peres} : Das Quartett

$$(r_{\gamma_D}, r_{\gamma_E}, r_{\gamma'_D}, r_{\gamma'_E})$$

mit zwei tatsächlich gemessenen Werten und zusätzlich zwei nicht gemessenen Werten existiert genau so real wie das Dublett

$$(r_{\gamma_D}, r_{\gamma_E})$$

der beiden tatsächlich gemessenen Werte.

A2_{Peres} : Die Einstellung der Messgeräte (also die Winkel, auf die die Stern-Gerlach-Magnete bzw. Polarisatoren einjustiert werden), und die Eigenschaften der Teilchen, die gemessen werden sollen, werden *nicht* durch eine gemeinsame Ursache vorherbestimmt. (Kein „Super-Determinismus“)

A3_{Peres} : Das zukünftige Ergebnis einer Messung beeinflusst *nicht* die Einstellung der Messgeräte (also die Winkel, auf die die Stern-Gerlach-Magnete bzw. Polarisatoren einjustiert werden). (Keine „Rückwärts-Verursachung“)

die Bell'sche Ungleichung (6.16)

$$-2 \leq \bar{q} \leq +2$$

$$\text{mit } \bar{q} = \overline{r_{\gamma_D} \cdot r_{\gamma_E}} + \overline{r_{\gamma_D} \cdot r_{\gamma'_E}} + \overline{r_{\gamma'_D} \cdot r_{\gamma_E}} - \overline{r_{\gamma'_D} \cdot r_{\gamma'_E}}$$

hergeleitet. Diese Ungleichung wird in den geschilderten Experimenten verletzt. Also muss mindestens eine der drei grundlegenden Annahmen $A1_{\text{Peres}}$, $A2_{\text{Peres}}$, $A3_{\text{Peres}}$ falsch sein. Die Annahmen $A2_{\text{Peres}}$ und $A3_{\text{Peres}}$ sind zwar nicht absolut sicher richtig, aber doch in hohem Maße wahrscheinlich. Also muss $A1_{\text{Peres}}$ die falsche Annahme sein. Demnach ergibt sich aus der Verletzung der Bell'schen Ungleichung, dass die Negation von $A1_{\text{Peres}}$ gilt:

- * Die Verletzung der Bell'schen Ungleichung beweist⁵⁶, dass die nicht gemessenen Werte $r_{\gamma'_D}$ und $r_{\gamma'_E}$ *nicht* genau so real existieren wie die tatsächlich gemessenen Werte r_{γ_D} und r_{γ_E} .

Diese Schlussfolgerung kann man auch als Negation der Annahme **B** formulieren:

- * Die Verletzung der Bell'schen Ungleichung beweist⁵⁶, dass die Teilsysteme eines verschränkten Quantensystems die Eigenschaften, die später gemessen werden, *nicht* bereits vor der Messung haben. Bei der Messung werden im allgemeinen *nicht* bereits vorher existierende Eigenschaften festgestellt, sondern diese Eigenschaften werden erst durch die Messung *erzeugt*.

Die Einschränkung „im allgemeinen“ bezieht sich auf den Spezialfall, dass das untersuchte Quantensystem schon vor der Messung in einem Eigenzustand des Messgeräts präpariert wurde. Nur in diesem Spezialfall hat das Quantenobjekt bereits vor der Messung die Eigenschaft, die bei der Messung festgestellt wird. In allen anderen Fällen wird⁵⁶ die Eigenschaft erst bei der Messung erzeugt.

Wenn bestimmte Eigenschaften von verschränkten Quantensystemen vor der Messung nicht existieren, dann existieren diese Eigenschaften in nicht verschränkten Quantensystemen vor der Messung erst recht nicht. Man kann die Schlussfolgerung also etwas allgemeiner formulieren:

- * Die Verletzung der Bell'schen Ungleichung beweist⁵⁶, dass viele Eigenschaften von Quantenobjekten, wie eine bestimmte Polarisation, ein bestimmtes magnetisches Moment in Richtung einer bestimmten Raumachse, ein bestimmter Ort, ein bestimmter Impuls, ... , durch die Messung nicht festgestellt sondern *erschaffen* werden.

Außerdem gilt folgende Schlussfolgerung:

- * Die Verletzung der Bell'schen Ungleichung beweist⁵⁶, dass die Natur nicht-lokal handelt. Als im Experiment von Weihs et al. die Polarisation eines der beiden SPDC-erzeugten Photonen durch Messung erschaffen wurde, da wurde zugleich die Polarisation des 400 m entfernten Partner-Photons erschaffen. Und zwar genau so erschaffen, dass die richtige Korrelation zwischen den Polarisationen der beiden Photonen realisiert wurde.

Wahrscheinlich können nur Leser, die eine Ausbildung in Klassischer Physik genossen haben, ermesen, wie sensationell dies Ergebnis ist.

Eine weitere Schlussfolgerung – und zwar nach meiner Ansicht die allerwichtigste – setzt eine zusätzliche Annahme voraus:

A4_{Peres}: Messungen haben eindeutige Ergebnisse.

Diese Annahme dürfte die meisten Leser verblüffen. Die Annahme besagt zum Beispiel: Ein Silberatom wird vom Stern-Gerlach-Magneten entweder eindeutig zum Nordpol oder eindeutig zum

Südpol abgelenkt, aber nicht zweideutig zum Nordpol *und* zum Südpol. Ist das nicht selbstverständlich?

Tatsächlich hat die Suche nach plausiblen Erklärungen für die seltsamen Quantenphänomene einige Physiker dazu bewogen, sogar die selbstverständlich erscheinende Annahme $A4_{\text{Peres}}$ in Zweifel zu ziehen. In der „Viele-Welten-Interpretation“ wird angenommen, dass jedes einzelne Silberatom tatsächlich vom Magneten zum Nordpol *und* zum Südpol abgelenkt wird. Darüber werde ich im Abschnitt 9.3 berichten. Die Annahme $A4_{\text{Peres}}$ ist also nicht trivial. Wenn sie richtig ist, dann kann man Folgendes überlegen:

Wenn das Messergebnis vor der Messung noch nicht feststeht, sondern erst im Moment der Messung erschaffen wird, dann kann es vor der Messung auch nicht berechnet werden. Dann wird das Ergebnis der Messung nicht durch irgendein Naturgesetz bestimmt, das die Physiker bisher nicht kennen, aber vielleicht in Zukunft eines Tages entdecken werden. Dann steht vielmehr definitiv fest dass es so ein Naturgesetz überhaupt nicht gibt, dass die Natur sich also im Moment der Messung *irrational* für eines der möglichen Messergebnisse entscheidet – wenn sie sich denn nach Annahme $A4_{\text{Peres}}$ tatsächlich für ein eindeutiges Ergebnis entscheidet:

- * Wenn die Annahmen $A2_{\text{Peres}}$, $A3_{\text{Peres}}$, $A4_{\text{Peres}}$ alle drei richtig sind, dann beweist die Verletzung der Bell'schen Ungleichung dass das Ergebnis der Messung nicht durch irgendein bisher unbekanntes Naturgesetz bestimmt wird, sondern dass die Erschaffung des Messergebnisses im Moment der Messung tatsächlich ein *irrationaler* Akt der Natur ist, der der wissenschaftlichen Analyse für immer entzogen sein wird, dass bei der Erschaffung des Messergebnisses echter Zufall am Werk ist.

Einstein hat die Entdeckung der Bell'schen Ungleichung, und ihre Verletzung in den Experimenten, nicht mehr erlebt; er starb 1955.

Also werden wir nie erfahren, ob er angesichts dieser Entwicklung seine Meinung geändert hätte. Sein Biograph Abraham Pais (1918–2000) berichtete viele Jahre später über Gespräche, die er mit Einstein über dessen Einwände gegen die Quantentheorie geführt hatte: “We often discussed his notions on objective reality. I recall that during one walk Einstein suddenly stopped, turned to me and asked whether I really believed that the moon exists only when I look at it.” [55, page 907]

Selbstverständlich wusste Einstein, dass ein plötzliches Verschwinden des Mondes eine unübersehbare, gigantische Flutkatastrophe an allen Küsten der Erde auslösen würde. Das „Hinschauen zum Mond“ war nur eine plakative Kurzformel für das Unbehagen, mit dem er die Entwicklung betrachtete, die die Quantentheorie in den letzten dreißig Jahren seines Lebens genommen hatte. Solange er lebte hielt er unbeirrbar an der Überzeugung fest, dass die Realität so ist wie sie ist, dass sie eben nicht „durch Messung erschaffen“ wird.

Diese Überzeugung teilte Albert Einstein (1879–1955) mit seinem älteren Kollegen Max Planck (1858–1947), dem im Jahr 1900 mit der Entdeckung des Wirkungsquantums $h = (3.4)$ und seiner Formel (3.3) für das Spektrum der Schwarzen Strahlung der allererste Schritt in Richtung Quantentheorie gelungen war. Planck verfasste 1945 eine kurz gefasste „Wissenschaftliche Selbstbiographie“ [56], in der er – sicherlich auch unter dem Eindruck der Diskussionen um die Quantentheorie – schrieb: „Was mich zu meiner Wissenschaft führte und von Jugend auf für sie begeisterte, ist die durchaus nicht selbstverständliche Tatsache, daß unsere Denkgesetze übereinstimmen mit den Gesetzmäßigkeiten im Ablauf der Eindrücke, die wir von der Außenwelt empfangen, daß es also dem Menschen möglich ist, durch reines Denken Aufschlüsse über jene Gesetzmäßigkeiten zu gewinnen. Dabei ist von wesentlicher Bedeutung, daß die Außenwelt etwas von uns Unabhängiges,

Absolute darstellt, dem wir gegenüberstehen, und das Suchen nach den Gesetzen, die für dieses Absolute gelten, erschien mir als die schönste wissenschaftliche Lebensaufgabe.“ Man glaubt fast eine Selbstironie herauszuhören, wenn Planck einige Seiten später schreibt: „Eine neue wissenschaftliche Wahrheit pflegt sich nicht in der Weise durchzusetzen, daß ihre Gegner überzeugt werden und sich als belehrt erklären, sondern vielmehr dadurch, daß ihre Gegner allmählich aussterben und daß die heranwachsende Generation von vornherein mit der Wahrheit vertraut gemacht ist.“ [56, Seite 22]

8 Which Way?

Single-particle-interference is based on the fact that the particle's trajectory is delocalized across both slits of a double-slit experiment or across both arms of an interferometer. If the trajectory is confined to a single slit or a single arm of the interferometer, the interference disappears. Experiments that investigate how the restriction of the trajectory occurs in detail are in the jargon of physicists referred to as “which-way” experiments. Some particularly noteworthy “which-way” experiments are described in this chapter. Upfront, I will discuss an important theoretical work on this topic, which Heisenberg published in spring 1927.

8.1 Heisenberg's Indeterminacy-Relation

The magnetic moment along a specific spatial axis (or the polarization in the case of photons) is an example of a relational property that the quantum object does not simply “have”, but rather one that is created through the interaction of this object with appropriate measuring instruments. This is an extreme example in that — as explained in fig. 6.2 — a magnetic moment previously created by measurement along a specific spatial axis is completely annihilated as soon as a further measurement creates the magnetic moment along another spatial axis. This is, so to speak, black-and-white painting: The magnetic moment in the direction of one axis is precisely defined, while it does not exist at all in the direction of all other axes.

There are other properties of quantum objects that are also

incompatible with one another, but which can be created in more nuanced gradations. Particularly important properties of this kind are position and momentum. In the diagrams 3.10 on page 69 and 4.14 on page 97, I have attempted to illustrate how the design and arrangement of the measuring apparatus determines the extent of the position of a photon or atom. The same applies to momentum: the accuracy with which the momentum is determined depends on the precision of the measuring instruments, through whose interaction with the quantum object the momentum of that object is created.

We will soon see that compromises must be made if one wishes to simultaneously determine the position and momentum of a quantum object using a measuring device: the more precisely the device determines the position, the less precisely it determines the momentum, and vice versa. We denote the uncertainty with which a property of an object is determined by measurement with the Greek letter Δ (pronounced: delta). For example, if the position of an atom in the x -direction is determined by passing it through a slit $1\ \mu\text{m}$ wide, then the uncertainty of its position in this direction is $\Delta x = 1\ \mu\text{m}$. We denote the uncertainty of the object's momentum in the x -direction, which is also determined by the precision of the measuring apparatus, as Δp_x .

In spring 1927, Heisenberg [36] showed that from the mysterious formalism of quantum theory the indeterminacy relation

$$\left. \begin{aligned} \Delta x \cdot \Delta p_x &\approx h \\ \Delta y \cdot \Delta p_y &\approx h \\ \Delta z \cdot \Delta p_z &\approx h \end{aligned} \right\} = 6.63 \cdot 10^{-34} \frac{\text{kg m}^2}{\text{s}} \quad (8.1)$$

can be derived,⁶⁸ where h is the Planck constant. Considering how tiny this constant is, it immediately becomes clear why the undeter-

⁶⁸ for physicists: Shortly thereafter, Robertson derived instead of (8.1) the more precise form $\Delta j \cdot \Delta p_j \approx \hbar/2$, see [57].

minacies of position and **momentum** went unnoticed until the end of the 19th century: The momentum of an object is proportional to its mass, and the masses of stones or even planets — and yet of the tiniest dust particles that could be detected microscopically at that time — are so enormous that the undeterminacy relation (8.1) could impossibly have been observed. It was only when physicists were able to examine individual atoms, electrons, and neutrons that the undeterminacies became apparent.

Heisenberg interpreted the undeterminacy relation (8.1) as follows: If the apparatus that simultaneously determines the position and momentum of an object is constructed such, that the momentum in x -direction is created with a fairly precise value (Δp_x very small), then this apparatus will — because $\Delta x \approx h/(\Delta p_x)$ is very large — create the object's position in x -direction only very imprecisely determined (i. e. the position in x -direction will be a widely extended region of space), and vice versa. It is impossible to construct the apparatus such that Δx and Δp_x are both arbitrarily small at the same time. The same applies to the y - and z -direction.

Why, for example, the position and momentum of an electron cannot be determined with arbitrary precision at the same time becomes clear from fig. 8.1 on the following page: The electron, whose position in x -direction is initially only very roughly determined, flies — as indicated by the turquoise arrows — from the left toward a slit with width Δx . When it passes through the slit (it could, of course, also miss the slit, but we are not interested in that case), then Δx is the uncertainty of its position in x -direction.

Now we must remember that the electron has not only the properties of a particle but also those of a wave. Its wavelength λ is related to its momentum p by de Broglie's relation $\lambda \stackrel{(4.2b)}{=} h/p$. When the electron wave passes through the slit, it is **diffracted**. At an angle of α , all partial waves passing through the slit at a distance of $\Delta x/2$ interfere destructively in pairs (i. e. they cancel

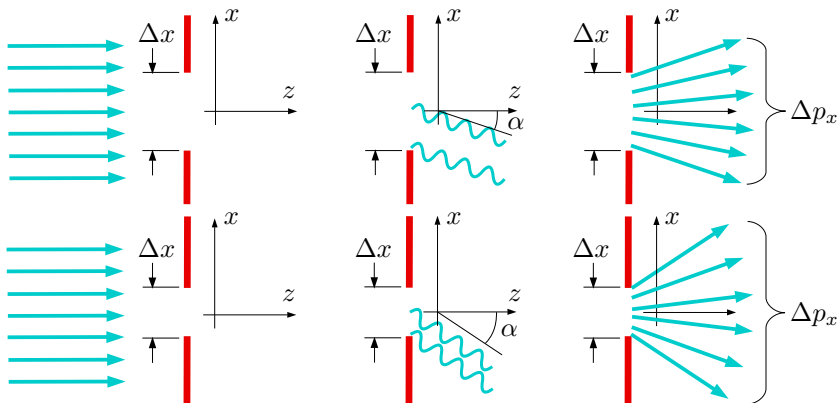


Fig. 8.1: The product $\Delta x \cdot \Delta p_x$ with Δx large (upper sketches) and Δx small (lower sketches)

each other out in pairs) when they superimpose at a large distance from the slit.

α is the smallest angle at which an interference minimum occurs. Although there are additional secondary maxima at larger diffraction angles, we obtain a useful rule of thumb by assuming for simplicity that the first minimum limits the electron's momentum in x -direction, as shown in the two sketches on the right in fig. 8.1.

If one wishes to reduce the undeterminacy Δx of an electron's position in x -direction by narrowing the slit, then due to

$$\Delta p_x \sim \alpha \stackrel{(4.4a)}{\sim} \frac{1}{\Delta x/2}$$

its momentum will inevitably be created with greater undeterminacy Δp_x . The smaller Δx is chosen, the larger the angle α of the first diffraction minimum becomes, as can be seen by comparing the lower sketches with the upper sketches in fig. 8.1. This is quantitatively formulated in Heisenberg's undeterminacy relation (8.1).

Thus the — in the classical understanding of these terms — mutually incompatible properties of waves and particles are preventing a quantum object from having both a precisely defined momentum and a precisely defined position at the same time. Bohr described such properties as *complementary*.

The dual nature of quantum objects as waves and particles — which gives rise to the indeterminacy relation (8.1) discovered by Heisenberg — is also important when we consider how, for example in the double-slit experiment with electrons, the trajectory of the particles can be restricted so that each electron passes through only one of the two slits, but not through both slits simultaneously. One could e. g. place behind one slit a vertically directed beam of light that is so intense that the electron will scatter at least one photon if it uses that slit. If a scattered photon is detected, then the electron must have passed through that slit. And it must have passed through the other slit if no scattered photon is detected. The method only works if the electron deflects at least one photon so strongly that it is actually clearly kicked out of the light beam. The electron must therefore transfer a considerable amount of momentum to the photon. Due to conservation of momentum, that kick will as well change the electron's momentum, and consequently also its

$$\text{de Broglie-wavelength} = \lambda \stackrel{(4.2b)}{=} \frac{h}{p} = \frac{h}{\text{momentum}} .$$

If the electron's de Broglie-wavelength is altered and disrupted to varying degrees along different paths, then interference is, of course, no longer possible. For decades, it seemed as though this fully explained the connection between the creation of a precisely defined trajectory and the disappearance of single-particle interference. That is why experiments, in which a precisely defined trajectory of particles could be created without significantly altering their de Broglie-wavelength, attracted great attention among physicists.

Successful experiments of that type were first conducted in the late 1990s.

8.2 The Path through the Interferometer

In 1998, S. Dürr, T. Nonn, and G. Rempe published the results of interference experiments with rubidium atoms that they had conducted at the University of Konstanz [58]. The setup of their experiment, outlined in fig. 8.2, resembles that of the interference experiment with neon atoms by Shimizu et al., which was sketched in fig. 4.12 on page 94.

In an evacuated chamber, a cloud of rubidium atoms was cooled to approximately $10 \mu\text{K}$ (10 millionths of a degree above absolute zero) using laser cooling. When the lasers were turned off, the atoms fell like stones, first through a $100 \mu\text{m}$ -wide slit S1 located 1 cm below the cloud, and then through a $450 \mu\text{m}$ -wide slit S2 located 20 cm below the cloud. Another 25 cm deeper, the x -position of the atoms was detected with a resolution of $50 \mu\text{m}$.

The slits were far too wide to observe diffraction effects from the atoms; their sole purpose was to limit the width of the trajectory.

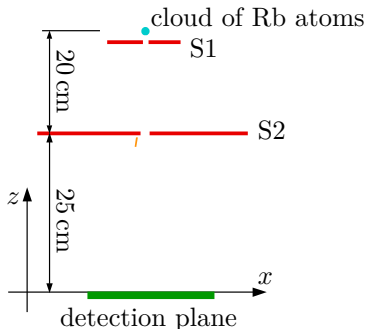


Fig. 8.2: Rubidium-atoms fall down in an evacuated chamber

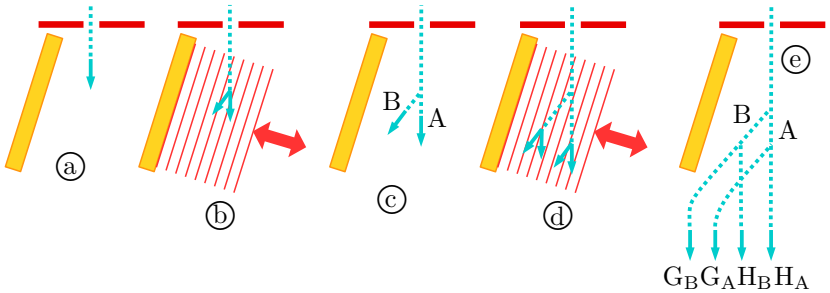


Fig. 8.3: The Rb-Interferometer with two Beam Splitters

Immediately after passing through slit S2, the atoms were irradiated with the standing wave field of a laser. This means the following:

If a laser beam is reflected exactly back into itself, regions (known as nodal planes) where the light intensity is zero form at regular intervals in front of the mirror, while the light intensity is maximum in the middle between the nodal planes. In fig. 8.3, the mirror located immediately below the slit S2 is sketched in yellow, and the nodal planes of the laser beam are indicated by red lines. The distance from one nodal plane to the next is exactly half a wavelength of the laser light.

The turquoise arrows and dashed lines are intended to represent the path of the rubidium atoms. The regular stack of regions of maximum laser intensity acts like a beam splitter for the beam of rubidium atoms. The greater the intensity of the laser light and the longer the laser remains switched on, the greater the probability that a rubidium atom passing through the laser's standing wave field will be reflected. The experimenters left the laser on for $45 \mu\text{s}$ and adjusted the intensity of the laser beam so that each Rb-atom was reflected with a probability of $1/2$ and continued straight ahead with a probability of $1/2$, see fig. 8.3 (b).

After $45 \mu\text{s}$, the laser was turned off for $105 \mu\text{s}$. During these $105 \mu\text{s}$, each atom continued with probability $1/2$ on path A, but

continued with probability 1/2 on path B, see fig. 8.3(c). Then the laser was turned on a second time for $45 \mu\text{s}$, so that once again each atom with probability 1/2 was deflected, but with probability 1/2 continued straight ahead, see fig. 8.3(d). After the second period of $45 \mu\text{s}$, the laser was finally switched off, and the atoms continued in free fall.

Let's estimate⁶⁹ the velocity of the atoms during their interaction with the laser field. If a stone is dropped in Earth's gravitational field (gravitational acceleration 9.81 m/s^2), then after a fall of 20 cm it's

$$\text{velocity} = 9.81 \frac{\text{m}}{\text{s}^2} \sqrt{\frac{2 \cdot 20 \text{ cm}}{9.81 \text{ m/s}^2}} = 1.98 \frac{\text{m}}{\text{s}} \approx 2 \frac{\text{m}}{\text{s}} .$$

The interaction with the laser field took place over a period of $45 \mu\text{s}$ plus a $105 \mu\text{s}$ pause plus another $45 \mu\text{s}$, totaling $195 \mu\text{s} \approx 0.2 \text{ ms}$. During this time, the atoms fell down approximately

$$2 \frac{\text{m}}{\text{s}} \cdot 0.2 \text{ ms} = 0.4 \text{ mm} .$$

The laser field had a width of 8 mm [59] in z -direction (see fig. 8.2), so that one can be certain that every Rb-atom experienced both on-phases of the laser in full length.

It should be emphasized that the angles of reflection are shown extremely exaggerated in fig. 8.3. The deflection angles in fact were so small that the rays G_A and G_B , when detected 25 cm lower, were offset by only about 1 mm relative to the rays H_A and H_B . Furthermore, the two rays G_A and G_B and the two rays H_A and H_B were not exactly parallel, but overlapped in the detection plane, causing them to interfere with one another. The effect of this interference is shown in fig. 8.4 on the facing page by blue triangles.

⁶⁹ Readers unfamiliar with such computations should not be alarmed but simply take note of the result.

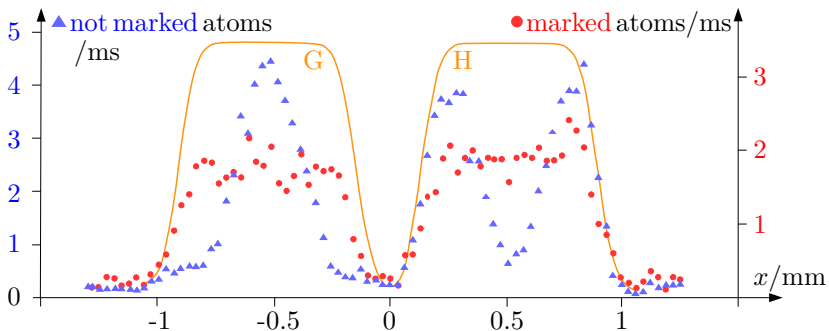


Fig. 8.4: Interference of not marked and marked Rb atoms

The solid yellow line indicates the region where the atoms of beams G_A and G_B , and beam H_A and H_B , respectively, reached the detection plane. A distinct maximum is visible at the center of the G beams, and a distinct minimum at the center of the H beams. This means that the Rb atoms in the beams G_A and G_B interfered constructively, whereas the beams H_A and H_B interfered destructively.

The explanation for this is *exactly* the same as for the photon-interference indicated in fig. 2.2 on page 24: At every beam splitter, there is a phase shift of a quarter wavelength between the transmitted and reflected waves. The reason for this phase shift is irrelevant to our considerations.¹² The two beams G_A and G_B were each reflected $1\times$, so they are “in phase” when they arrive at the detection plane. In contrast, the atoms in beam H_B were reflected twice, while the atoms in beam H_A were not reflected at all; see fig. 8.3©. Overall, the phase of the atoms in the H_B beam is thus shifted by half a wavelength relative to the phase of the atoms in the H_A beam, so that these beams interfere destructively as explained in fig. 2.4 on page 26.

This is true only at the center of regions G and H, however, where

the lengths of the paths of the two partial beams is exactly equal. At the right and left edges of the regions where the beams reach the detection plane, the path of one beam is half a wavelength longer than that of the other beam. Therefore, constructive interference is observed at the right and left edges of region H, but destructive interference is observed at the right and left edges of region G.

At every point in the detection plane, less than 5 atoms were recorded per millisecond. This means that the atoms arrived at each point at an average time interval of 0.2 ms or more. Their velocity in the detection plane was

$$\text{velocity} = 9.81 \frac{\text{m}}{\text{s}^2} \sqrt{\frac{2 \cdot 45 \text{ cm}}{9.81 \text{ m/s}^2}} \approx 3 \frac{\text{m}}{\text{s}} ,$$

so that the average distance from one atom to the next was more than

$$3 \frac{\text{m}}{\text{s}} \cdot 0.2 \text{ ms} = 0.6 \text{ mm} .$$

Two atoms separated by such a large distance certainly cannot interfere with each other. The interference therefore arose because the position of each individual atom was so imprecisely defined (i. e. so extended) that it covered multiple paths through the interferometer, and thus each individual atom interfered with itself.

As reported so far, the experiment by Dürr et al. was “merely” another (very beautiful) example of self-interference of atoms. The real highlight came in the second part of the experiment: Now the trajectory of the atoms, which had previously extended along paths A *and* B (see fig. 8.3 on page 207), was narrowed so that it extended only along path A *or* B. The more precisely defined trajectory was created by a path-dependent manner of marking the Rb atoms.

In its ground state, the rubidium atom has an angular momentum quantum number $j = 2$. The energy of the excited state

with angular momentum quantum number $j = 3$ is $h \cdot 3.04$ GHz higher. We define the following notation for the state vectors of the rubidium atoms:

$$|2\rangle = \text{ground state with } j = 2$$

$$|3\rangle = \text{excited state with } j = 3$$

To mark the atoms, those traveling along path A were excited to the $j = 3$ state by irradiation with 3.04 GHz, while the atoms on path B were not excited. Anyone interested in the details of the intricate marking procedure can find more information in appendix A.7. For our discussion, it suffices to know what the state vectors of marked and not-marked rubidium atoms were when they reached the detector plane:

$$\begin{aligned} |\text{Rb}\rangle_{\text{not marked}} &\stackrel{\text{(A.29c)}}{=} \\ &= \frac{1}{2} \left(-|2\rangle|G_A\rangle - |2\rangle|G_B\rangle + |2\rangle|H_A\rangle + |2\rangle|H_B\rangle \right) \end{aligned} \quad (8.3a)$$

$$\begin{aligned} |\text{Rb}\rangle_{\text{marked}} &\stackrel{\text{(A.28e)}}{=} \\ &= \frac{1}{2} \left(i|3\rangle|G_A\rangle - |2\rangle|G_B\rangle + i|3\rangle|H_A\rangle + |2\rangle|H_B\rangle \right) \end{aligned} \quad (8.3b)$$

Here the state vectors are written as products, where the first factor indicates whether the atom's angular momentum quantum number is $j = 2$ or $j = 3$, while the second factor indicates which of the four possible paths the atom has taken.

In both $|\text{Rb}\rangle$ state vectors (8.3), the vectors $|G_A\rangle$, $|G_B\rangle$, $|H_A\rangle$, $|H_B\rangle$ of the four possible paths are entangled. In (8.3b) there is a clear correlation, however, between $|3\rangle$ and path A, as well as between $|2\rangle$ and path B. In (8.3b), the marking by the state $|3\rangle$ has restricted the atom's trajectory to path A, and the marking by the state $|2\rangle$ has restricted its trajectory to path B. The restriction of the trajectory to path A or B due to the marking of the atom is

an objective fact. A measurement could unambiguously determine whether an atom is in the state $|3\rangle$ or $|2\rangle$, and one would then know whether it has taken path A or path B.

In the detection plane, the x -axis was scanned over a length of 2.6 mm with a step size of 2.6 mm/80, and at each of the 81 measurement positions the number of incoming rubidium atoms was counted (the count rates plotted in diagram 8.4 are the averages of numerous runs of the experiment). If an atom is detected at position x_j , then the detector prepares it in its eigenstate $|x_j\rangle$, with $j = 1 \dots 81$. The probability that an atom is detected at position x_j is, according to Born's rule (5.19b), equal to the square of the projection amplitude of $|x_j\rangle$ onto $|(8.3)\rangle$.

To keep it simple, we calculate the projection amplitude for just one point far to the left in the diagram 8.4, for example at x_{17} . Then we can be fairly certain that only atoms that have taken the paths G_A or G_B arrive there, but no atoms from the paths H_A or H_B . With this simplification, the probability of detecting a *not-marked* atom at position x_{17} is equal to

$$\begin{aligned}
 P(x_{17})_{\text{not marked}} &\stackrel{(8.3a)}{=} \frac{1}{4} \underbrace{|\langle 2|2\rangle|^2}_{=1} \left| \langle x_{17} | G_A \rangle + \langle x_{17} | G_B \rangle \right|^2 = \\
 &\stackrel{(5.5j), (5.6c)}{=} \frac{1}{4} \left(\underbrace{|\langle x_{17} | G_A \rangle|^2}_{\textcircled{1}} + \underbrace{\langle G_A | x_{17} \rangle \langle x_{17} | G_B \rangle}_{\textcircled{2}} + \right. \\
 &\quad \left. + \underbrace{\langle G_B | x_{17} \rangle \langle x_{17} | G_A \rangle}_{\textcircled{3}} + \underbrace{|\langle x_{17} | G_B \rangle|^2}_{\textcircled{4}} \right). \quad (8.4a)
 \end{aligned}$$

In contrast, because of

$$\langle 2|3\rangle = \langle 3|2\rangle = 0 \quad \text{and} \quad |\langle 2|2\rangle|^2 = |\langle 3|3\rangle|^2 = 1,$$

the probability to detect a *marked* atom at position x_{17} is equal to

$$\begin{aligned}
P(x_{17})_{\text{marked}} &\stackrel{(8.3b)}{=} \frac{1}{4} \left| i\langle 3 | \langle x_{17} | | G_A \rangle - | 2 \rangle \langle x_{17} | | G_B \rangle \right|^2 = \\
&\stackrel{(5.5j), (5.6c)}{=} \frac{1}{4} \left(-i \langle G_A | | x_{17} \rangle \langle 3 | - \langle G_B | | x_{17} \rangle \langle 2 | \right) \cdot \\
&\quad \cdot \left(i\langle 3 | \langle x_{17} | | G_A \rangle - | 2 \rangle \langle x_{17} | | G_B \rangle \right) = \\
&= \frac{1}{4} \left(\underbrace{\left| \langle x_{17} | | G_A \rangle \right|^2}_{\textcircled{1}} + \underbrace{\left| \langle x_{17} | | G_B \rangle \right|^2}_{\textcircled{4}} \right). \tag{8.4b}
\end{aligned}$$

In the terms ① and ④ in (8.4), the atom has reached position x_{17} clearly via path A or clearly via path B. Such terms do not describe interference. Only such terms appear in (8.4b).

In contrast, the terms ② and ③ in (8.4a) describe the case where the atom has taken both paths (i. e. its trajectory was delocalized over both paths). It are these two terms that cause the interference maxima and minima of the blue triangles in diagram 8.4.

The results of the experiment with marked rubidium atoms are plotted as red dots in diagram 8.4. As in any experiment, the result is not perfect. If one looks closely, one can still make out traces of interference peaks. Compared to the blue triangles, however, the interference has largely disappeared in case of the red dots.

The interference thus disappears due to the marking of the atoms dependent on their path (A or B). This is quite different from Heisenberg's considerations regarding the mutual disturbance of measurements of position and momentum (and thus the de Broglie-wavelength) of an object, which underlie his indeterminacy relation (8.1). In the experiment by Dürr et al. the de Broglie-wavelength of the rubidium atoms was virtually not disturbed at all by the creation of the precisely defined trajectory of the atoms. The tiny energy transfer

$$E \stackrel{(4.1a)}{=} h\nu = h \cdot 3.04 \text{ GHz}$$

and the tiny momentum transfer

$$p \stackrel{(4.1b)}{=} \frac{E}{c} = \frac{h \cdot 3.04 \text{ GHz}}{c}$$

due to the marking of the atoms by means of the microwave pulses is almost negligible. Nevertheless, the interference disappeared. So the interference does not disappear because of a disturbance of the de Broglie-wavelength or any other disturbance affecting the atoms, but simply because due to marking of the atoms a more precisely defined trajectory is created than without marking.

The trajectory created by the two slits (see fig. 8.2) extends across both paths A and B, just as the trajectory (i. e. the temporal sequence of positions) of the neon-atom in figure 4.14 on page 97 extends across both slits of the double slit. By marking the rubidium atom, however, its trajectory is restricted to one of the paths A *or* B.

The creation of the narrower trajectory is an objective process that eliminates interference. In their experiment, Dürr et al. did not investigate whether the marked atoms were in state $|2\rangle$ or in state $|3\rangle$. The objective fact that the marking restricted the atom's trajectory to one of the paths A *or* B was sufficient to cause the self-interference to disappear.

8.3 A Quantum-Eraser

Using the setup outlined in fig. 8.5 on the next page, Kim, Yu, Kulik, Shih, and Scully [60] demonstrated in 1998 the possibility of a “quantum-eraser”. In the sequel I will explain what that is.

Using SPDC⁷⁰, pairs of photons were generated in an entangled state as follows: A BBO-crystal (β -barium-borate BaB_2O_4) is

⁷⁰ SPDC-type I was explained in section 3.5, SPDC-type II in section 7.3. Kim et al. used SPDC-type II, but this detail was irrelevant to their experiment.

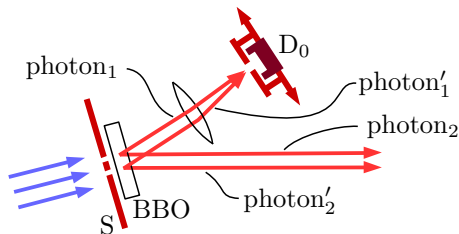


Fig. 8.5: Interference?

irradiated from the left with a pump laser (indicated by the blue arrows, wavelength 351.1 nm). Immediately in front of the crystal is a metal-plate with a

double-slit S : (8.5)

width = 0.3 mm , distance center-to-center = 0.7 mm

Most ultraviolet pump photons pass through the BBO crystal unchanged. Only occasionally a 351.1 nm pump photon is converted into two 702.2 nm photons, which leave the crystal at an angle of approximately 3° (relative to the pump beam). We call the photon pairs generated behind the upper slit photon_1 and photon_2 , and the photon pairs generated behind the lower slit photon'_1 and photon'_2 .

photon_1 and photon'_1 pass through a convex lens whose focal plane is raster-scanned with detector D_0 . Can we expect to observe an interference pattern in this case? The conversion of a pump photon into a pair of 702.2 nm photons occurs so rarely that practically never two pairs of daughter photons are generated within a few nanoseconds. Therefore, practically never two photons arrive at the detector D_0 simultaneously. But by now we have seen enough examples of self-interference of individual particles. If the trajectory of the pump photons is so imprecisely localized that it covers both slits, could a photon be generated in a delocalized

state behind the upper and lower slits, and interfere with itself in the focal plane of the lens?

As we will see in a moment, this is indeed possible, provided that the point of origin of the photon — behind the upper or the lower slit? — is truly not localized. This is exactly as in the conventional double-slit experiment: If the particle's trajectory through the double slit is localized only so imprecisely that it can take the path through both slits, then it will interfere with itself. But if — as in the example of the marked rubidium atoms — the particle's path is localized so precisely that it uses only one of the two slits, then there is no interference.

In the experiment shown in fig. 8.5, the point of origin of the photon is confined to the region behind the upper or the lower slit — unless additional measures are taken. This is because the partner photon is generated at the same time. The location of the SPDC process is determined by the partner photon. Why? This is explained in fig. 8.6. To separate the paths of photon_2 and photon'_2 , a prism was inserted between the crystal and the detectors D_3 and D_4 . If the partner photon is detected by detector D_4 , then the photon pair was generated behind the upper slit. If it is detected by detector D_3 , then the photon pair was generated behind the lower slit.

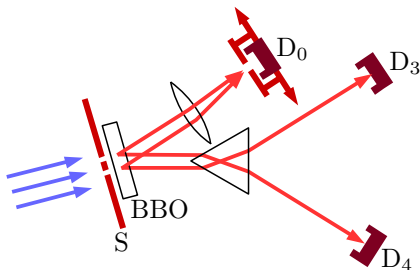


Fig. 8.6: The origin of the photon is being localized.

The point of origin of photon₁ resp. photon'₁ is created by the measurement apparatus. The measurement apparatus, that is in this case the BBO crystal, the double-slit, the beam of pump photons, and the outgoing photon₂ or photon'₂. In the SPDC-process, this apparatus configuration creates not only the pair of daughter photons but also their point of origin, located behind the upper or lower slit. The creation of the point of origin is an objective fact. It is not necessary for an experimenter to take note of the point of origin of the daughter photons using the setup shown in fig. 8.6. Even if we ignore photon₂ or photon'₂, or let them simply hit a wall, thereby destroying the information about their point of origin, there is no self-interference of photon₁ or photon'₁.

Let's first double-check that this fact is correctly described by quantum theory: In the SPDC-process, when a pump photon with wavelength 351.1 nm passes through the upper slit and is converted into two photons with wavelength 702.2 nm, the pair of daughter photons emerges behind this slit in state $|1\rangle|2\rangle$. If the pump photon passes through the lower slit and is converted into two photons with wavelength 702.2 nm, then the pair of daughter photons emerges behind this slit in state $|1'\rangle|2'\rangle$. If the trajectory of the pump photon is delocalized across both slits, however, then the pair of daughter photons emerges in the entangled state

$$|\text{photon-pair}\rangle = \frac{1}{\sqrt{2}} \left(|1\rangle|2\rangle + |1'\rangle|2'\rangle \right). \quad (8.6)$$

In the setup shown in fig. 8.5 on page 215, there is only one detector, namely D_0 . This detector scans 28 different x -positions in the focal plane of the lens. Consequently

$$|x_{D_0}\rangle \quad (8.7)$$

with 28 different x -values are the 28 eigenvectors of the detector D_0 . According to Born's rule (5.19b), the probability of a specific

measurement result is equal to the modulus-square of the projection amplitude of (8.6) onto the corresponding eigenvector of the detector. Thus, the probability $P(x)$ that D_0 detects a photon at position x is equal to

$$\begin{aligned}
 P(x) &= \frac{1}{2} \left| \langle x_{D0} || 1 \rangle |2\rangle + \langle x_{D0} || 1' \rangle |2'\rangle \right|^2 = \\
 &\stackrel{(5.5j), (5.6c)}{=} \frac{1}{2} \left(\langle 2 || 1 \rangle \langle x_{D0} \rangle + \langle 2' || 1' \rangle \langle x_{D0} \rangle \right) \cdot \\
 &\quad \cdot \left(\langle x_{D0} || 1 \rangle |2\rangle + \langle x_{D0} || 1' \rangle |2'\rangle \right) = \\
 &= \frac{1}{2} \left(\underbrace{\langle 2 || 2 \rangle}_1 \underbrace{\langle 1 || x_{D0} \rangle \langle x_{D0} || 1 \rangle}_{\textcircled{1}} + \underbrace{\langle 2 || 2' \rangle}_0 \underbrace{\langle 1 || x_{D0} \rangle \langle x_{D0} || 1' \rangle}_{\textcircled{2}} \right. \\
 &\quad \left. + \underbrace{\langle 2' || 2 \rangle}_0 \underbrace{\langle 1' || x_{D0} \rangle \langle x_{D0} || 1 \rangle}_{\textcircled{3}} + \underbrace{\langle 2' || 2' \rangle}_1 \underbrace{\langle 1' || x_{D0} \rangle \langle x_{D0} || 1' \rangle}_{\textcircled{4}} \right) = \\
 &= \frac{1}{2} \left(\underbrace{\left| \langle x_{D0} || 1 \rangle \right|^2}_{\textcircled{1}} + \underbrace{\left| \langle x_{D0} || 1' \rangle \right|^2}_{\textcircled{4}} \right). \tag{8.8}
 \end{aligned}$$

Due to $\langle 2 || 2' \rangle = \langle 2' || 2 \rangle = 0$, the “mixed” terms $\textcircled{2}$ and $\textcircled{3}$, in which both 1 and 1' appear, have been lost. But precisely these are the two terms through which interference could have occurred. In $\textcircled{1}$, only photon 1 appears, which was unambiguously generated behind the upper slit. And in $\textcircled{4}$, only photon 1' appears, which was unambiguously generated behind the lower slit. These two terms cannot contribute to interference.

For comparison, let us consider a simpler experiment, which is shown in the left-hand sketch of fig. 8.7 on the facing page. (The right-hand side of fig. 8.7 simply reprints fig. 8.5 for comparison.) In the simpler experiment, the BBO-crystal is removed, and a beam of photons with wavelength 702.2 nm is directed through the double slit onto the detector D_0 . The beam is assumed to be

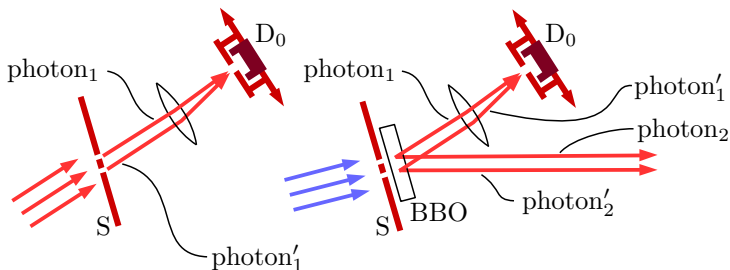


Fig. 8.7: Unprecise (left sketch) and precise (right sketch) localized trajectory of photon₁ resp. photon'₁

uniformly delocalized across both slits, so that the state vector of the photons in the focal plane of the lens is

$$\frac{1}{\sqrt{2}} \left(|1\rangle + |1'\rangle \right). \quad (8.9)$$

A photon₂ or photon'₂ no longer exists in this setup. The probability $P(x)$ that the photon is observed by detector D_0 at position x is, according to Born's rule (5.19b), equal to the modulus-square of the projection amplitude of (8.9) onto the corresponding eigenvector of the detector.

$$\begin{aligned}
 P(x) &= \frac{1}{2} \left| \langle x_{D0} | \left(|1\rangle + |1'\rangle \right) \right|^2 = \\
 &\stackrel{(5.5j), (5.6c)}{=} \frac{1}{2} \left(\langle 1 || x_{D0} \rangle + \langle 1' || x_{D0} \rangle \right) \left(\langle x_{D0} || 1 \rangle + \langle x_{D0} || 1' \rangle \right) = \\
 &= \frac{1}{2} \left(\underbrace{\left| \langle x_{D0} || 1 \rangle \right|^2}_{\textcircled{1}} + \underbrace{\langle 1 || x_{D0} \rangle \langle x_{D0} || 1' \rangle}_{\textcircled{2}} + \right. \\
 &\quad \left. + \underbrace{\langle 1' || x_{D0} \rangle \langle x_{D0} || 1 \rangle}_{\textcircled{3}} + \underbrace{\left| \langle x_{D0} || 1' \rangle \right|^2}_{\textcircled{4}} \right). \quad (8.10)
 \end{aligned}$$

Different from (8.8), the terms ② and ③ now also are included in the final result. It are precisely these “mixed” terms, in which both 1 and 1' appear, that describe the self-interference of the photon. The value of these terms depends on the difference in path lengths from the upper or lower slit to the position x of the detector D_0 . If the difference in path lengths is equal to

$$0, \pm\lambda, \pm2\lambda, \pm3\lambda, \dots,$$

where $\lambda = 702.2 \text{ nm}$ is the wavelength of the photon, then there is an interference-maximum at this point. But if the difference in path lengths is equal to

$$\pm\lambda/2, \pm3\lambda/2, \pm5\lambda/2, \dots,$$

then there is an interference minimum at this point, see fig. 2.4 on page 26.

Thus interference seems impossible when photons are generated via SPDC. Using a “quantum eraser”, however, it is nonetheless possible to achieve self-interference even with these photons. Marlan Scully (together with Kai Drühl) had the idea for a quantum eraser as early as 1982 [61], but it was not until seventeen years later that he (together with Kim, Yu, Kulik, and Shih) succeeded in realizing it through the experiment described here.

The experimenters directed photon₂ and photon'₂ through the mirrors shown in yellow in fig. 8.8 on the facing page onto a beam splitter. This was not a polarizing beam splitter, but an ordinary one, which transmits each photon with probability 1/2 and reflects it with probability 1/2, regardless of its polarization. Behind the beam splitter, the photons were detected by detectors D₁ and D₂.

If D₁ or D₂ detects a photon, then, because of the beam splitter, it is impossible to know whether it was a photon₂ or a photon'₂. The beam splitter has erased this information, and with it, the information about the origin of the partner photon detected by D₀. Consequently, nothing stands in the way of the interference of photons at the detector D₀.

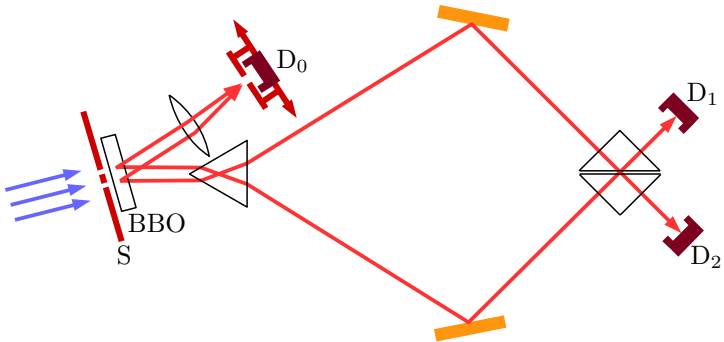


Fig. 8.8: A quantum-eraser

Well, somehow that doesn't sound very convincing. Didn't I just say that localization is an objective fact, regardless of whether or not anyone takes note of it? Why is the localization of the photons' point of origin erased by the beam splitter, but not erased when we let photon_2 or photon'_2 run unregistered against the wall?

Before we try to clarify this, let's first take a look at what exactly was observed in this experiment. First it must be noted that the experimental setup actually was slightly more complex than discussed so far, see fig. 8.9 on the next page. Additional beam splitters were inserted into the paths of both photon_2 and photon'_2 , which reflect the photons with probability $1/2$ to detectors D_3 and D_4 , respectively. When this happens, the situation is again the same as in the setup of fig. 8.6. If D_3 (resp. D_4) is triggered, then the origin of the photon pair is precisely localized in the BBO crystal behind the lower (resp. upper) slit, and consequently D_0 cannot observe any interference in the focal plane of the lens. If, on the other hand, D_1 or D_2 responds, then the origin of the photons is not localized, and therefore interference in the focal plane of the lens is possible.

In this setup, the choice of whether or not the origin of the

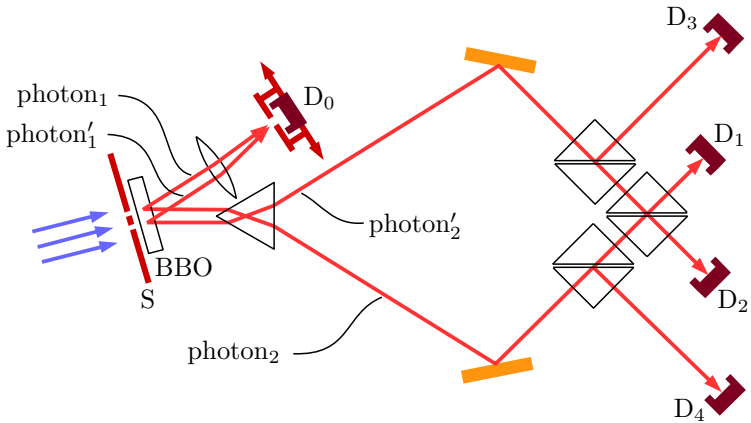


Fig. 8.9: A quantum-eraser with delayed choice

photons is localized is not made by the experimenters but by Nature herself, in that the photon is randomly (with probability $1/2$) reflected or randomly (with probability $1/2$) transmitted at the first beam splitter. The choice is referred to as “delayed” if it is made only after the detector D_0 has already registered the other photon of the entangled pair. This was the case in the experiment of Kim et al., because the first beam splitters were about 2.3 m farther from the BBO-crystal than the detector D_0 . To cover this distance difference, the photons need

$$\frac{2.3 \text{ m}}{\text{speed of light}} = \frac{2.3 \text{ m}}{3 \cdot 10^8 \text{ m/s}} \approx 7.7 \text{ ns} .$$

The response time of the detectors, on the other hand, was less than 1 ns. Thus, every photon had in fact already been detected by detector D_0 long before Nature made her choice (through reflection or transmission of the partner photon at the first beam splitter) whether or not the origin of the photons was localized.

The detectors D_1, D_2, D_3, D_4 were located another 20 cm away from the first beam splitters, so that the photons reached these

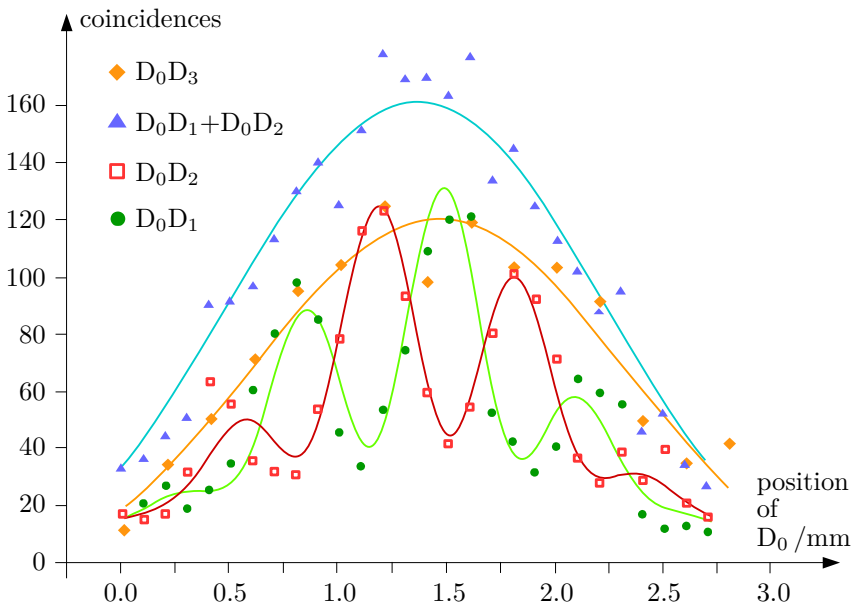


Fig. 8.10: Observed photon-pairs as function of D_0 position

detectors

$$\frac{2.5 \text{ m}}{3 \cdot 10^8 \text{ m/s}} \approx 8.3 \text{ ns}$$

later than their partner photons reached detector D_0 . Therefore, whenever detector D_0 was triggered, it was observed whether one of the detectors D_1, D_2, D_3, D_4 was triggered approximately 8 ns later. These coincidences were counted⁷¹, and plotted in diagram 8.10.

When measuring the coincidences between D_0 and D_1 (green

⁷¹ Kim et al. unfortunately missed to specify in their publication [60] the time interval over which the coincidences were counted, whether per millisecond, per minute, or ...

dots in the diagram) and between D_0 and D_2 (red squares in the diagram), the D_0 detector scanned the focal plane of the lens in steps of 0.1 mm; when measuring the coincidences of D_0 and D_3 (yellow diamonds in the diagram), the step size was 0.2 mm. The blue triangles are simply the sum of the green dots and the red squares.

The only result that is easy to understand are the yellow diamonds: If detector D_3 is triggered, then the origin of the photon pair is located at the lower slit. Thus, one expects that detector D_0 does not observe the interference pattern of a double slit, but rather the [diffraction](#) pattern of a single slit with width $\stackrel{(8.5)}{=} 0.3$ mm, which is shifted to the right by 0.35 mm $\stackrel{(8.5)}{=}$ (half slit distance) from the center of the overall image. This expectation is plotted as the yellow curve, and the yellow diamonds do indeed lie on this line – apart from the inevitable experimental fluctuations and inaccuracies. This line corresponds to the [yellow](#) term ④ in (8.8) resp. (8.10).

The other three curves, on the other hand, are somewhat surprising. If one of the detectors D_1 or D_2 is triggered, then the localization of the origin of the photon pair is supposedly “erased”. Shouldn’t detector D_0 observe an interference pattern in this case? The blue triangles, however, lie reasonably well on the blue line, which contains no interference structures whatsoever.

Only when the coincidences D_0D_1 (green dots) and D_0D_2 (red squares) are considered separately does a pattern emerge that resembles interference at a double slit. But, puzzlingly, the maxima and minima are shifted to the right and left. This can clearly be seen, for example, when comparing the diffraction of neutrons at a single slit and at a double slit in [fig. 4.6](#) on page 81. With “normal” diffraction, the central maximum lies exactly in the center of the diffraction pattern.

To understand [fig. 8.10](#), we need to take a closer look at the state

vectors and projection amplitudes of the photons. When photon₂' strikes the first beam splitter, it is transmitted or reflected toward detector D₃ with equal probability. Thus its state vector after the first beam splitter is

$$|2'\rangle \xrightarrow{\text{first beam-splitter}} \frac{1}{\sqrt{2}} \left(|2'_{\text{transmitted}}\rangle - i|2'_{\text{D}_3}\rangle \right). \quad (8.11)$$

Here, the notation $|2'_{\text{D}_3}\rangle$ is introduced for the state vector of photon₂' that strikes detector D₃. The factor $-i = -\sqrt{-1}$ arises because there is always a phase shift of one-quarter of a wavelength between the transmitted and reflected particles at a beam splitter.¹² According to (A.8), this phase shift is represented in theory by the factor $-i$. At the second beam splitter, the transmitted photon is transmitted to detector D₂ or reflected to detector D₁ with equal probability:

$$|2'_{\text{transmitted}}\rangle \xrightarrow{\text{second beam-splitter}} \frac{1}{\sqrt{2}} \left(|2'_{\text{D}_2}\rangle - i|2'_{\text{D}_1}\rangle \right) \quad (8.12)$$

Here again the phase-shift $-i$ of the reflected photon shows up. Inserting (8.12) into (8.11), we get

$$|2'\rangle \xrightarrow{\text{both beam-splitters}} \frac{1}{\sqrt{2}} \left\{ \frac{1}{\sqrt{2}} \left(|2'_{\text{D}_2}\rangle - i|2'_{\text{D}_1}\rangle \right) - i|2'_{\text{D}_3}\rangle \right\},$$

and by the same method we find

$$|2\rangle \xrightarrow{\text{both beam-splitters}} \frac{1}{\sqrt{2}} \left\{ \frac{1}{\sqrt{2}} \left(|2_{\text{D}_1}\rangle - i|2_{\text{D}_2}\rangle \right) - i|2_{\text{D}_4}\rangle \right\}.$$

This is inserted into the state vector of the photon-pair:

$$\begin{aligned} |\text{photon-pair}\rangle &\stackrel{(8.6)}{=} \frac{1}{\sqrt{2}} \left(|1\rangle|2\rangle + |1'\rangle|2'\rangle \right) = \\ &= \frac{1}{2} \left[\frac{1}{\sqrt{2}} |1\rangle|2_{\text{D}_1}\rangle - \frac{i}{\sqrt{2}} |1\rangle|2_{\text{D}_2}\rangle - i|1\rangle|2_{\text{D}_4}\rangle + \right. \\ &\quad \left. + \frac{1}{\sqrt{2}} |1'\rangle|2'_{\text{D}_2}\rangle - \frac{i}{\sqrt{2}} |1'\rangle|2'_{\text{D}_1}\rangle - i|1'\rangle|2'_{\text{D}_3}\rangle \right]. \quad (8.13) \end{aligned}$$

The eigenvectors of the detectors are

$$|x_{D_0}\rangle|D_1\rangle, |x_{D_0}\rangle|D_2\rangle, |x_{D_0}\rangle|D_3\rangle, |x_{D_0}\rangle|D_4\rangle, \quad (8.14)$$

because solely coincidences of D_0 and one of the four other detectors are registered. x_{D_0} again symbolizes the position of detector D_0 in the focal plane of the lense. Since 28 different x -positions were sampled, (8.14) stands for $4 \cdot 28 = 112$ different vectors. According to Born's rule (5.19b), the probability of a specific measurement result is equal to the modulus square of the projection amplitude of (8.13) onto the corresponding eigenvector of the detectors. For example, the probability that a photon is detected by D_0 at position x , and that about 8 ns later the partner photon is detected by D_3 , is equal to

$$\begin{aligned} & \left| \langle x_{D_0} | \langle D_3 | |8.13\rangle \right|^2 = \\ & = \left| \frac{1}{2} \left[\frac{1}{\sqrt{2}} \langle x_{D_0} ||1\rangle \underbrace{\langle D_3 ||2_{D_1}\rangle}_0 - \frac{i}{\sqrt{2}} \langle x_{D_0} ||1'\rangle \underbrace{\langle D_3 ||2'_{D_1}\rangle}_0 - \right. \right. \\ & \quad - \frac{i}{\sqrt{2}} \langle x_{D_0} ||1\rangle \underbrace{\langle D_3 ||2_{D_2}\rangle}_0 + \frac{1}{\sqrt{2}} \langle x_{D_0} ||1'\rangle \underbrace{\langle D_3 ||2'_{D_2}\rangle}_0 - \\ & \quad \left. \left. - i \langle x_{D_0} ||1'\rangle \underbrace{\langle D_3 ||2'_{D_3}\rangle}_1 - i \langle x_{D_0} ||1\rangle \underbrace{\langle D_3 ||2_{D_4}\rangle}_0 \right] \right|^2 = \\ & = \frac{1}{4} \underbrace{\left| \langle x_{D_0} ||1'\rangle \right|^2}_{\textcircled{4}}. \end{aligned} \quad (8.15a)$$

Here $|-i|^2 = 1$ has been used. This result gives the probability that a photon is created localized behind the lower slit and reaches detector D_0 at position x . (8.15a) is plotted as yellow curve in diagram 8.10. (8.15a) corresponds to the term $\textcircled{4}$ in (8.8) resp. (8.10). The different factors $1/4$ resp. $1/2$ stem from the fact that

the sum of the probabilities of the various possible outcomes in each experiment must equal 1 .

The probability that a photon is detected by D_0 at position x , and that about 8 ns later the partner photon is detected by D_4 , is equal to

$$\begin{aligned}
 & \left| \langle x_{D0} | \langle D_4 | | 8.13 \rangle \right|^2 = \\
 & = \left| \frac{1}{2} \left[\frac{1}{\sqrt{2}} \langle x_{D0} | | 1 \rangle \underbrace{\langle D_4 | | 2_{D1} \rangle}_0 - \frac{i}{\sqrt{2}} \langle x_{D0} | | 1' \rangle \underbrace{\langle D_4 | | 2'_{D1} \rangle}_0 - \right. \right. \\
 & \quad \left. - \frac{i}{\sqrt{2}} \langle x_{D0} | | 1 \rangle \underbrace{\langle D_4 | | 2_{D2} \rangle}_0 + \frac{1}{\sqrt{2}} \langle x_{D0} | | 1' \rangle \underbrace{\langle D_4 | | 2'_{D2} \rangle}_0 - \right. \\
 & \quad \left. - i \langle x_{D0} | | 1' \rangle \underbrace{\langle D_4 | | 2'_{D3} \rangle}_0 - i \langle x_{D0} | | 1 \rangle \underbrace{\langle D_4 | | 2_{D4} \rangle}_1 \right] \right|^2 = \\
 & = \frac{1}{4} \underbrace{\left| \langle x_{D0} | | 1 \rangle \right|^2}_{\textcircled{1}} . \tag{8.15b}
 \end{aligned}$$

Here again $|-i|^2 = 1$ has been used. This result gives the probability that a photon is created localized behind the upper slit and reaches detector D_0 at position x . (8.15b) corresponds to the term $\textcircled{1}$ in (8.8) resp. (8.10). The different factors 1/4 resp. 1/2 stem from the fact that the sum of the probabilities of the various possible outcomes in each experiment must equal 1 .

The probability that a photon is detected by D_0 at position x , and that about 8 ns later the partner photon is detected by D_1 , is equal to

$$\left| \langle x_{D0} | \langle D_1 | | 8.13 \rangle \right|^2 =$$

$$\begin{aligned}
&= \left| \frac{1}{2} \left[\frac{1}{\sqrt{2}} \langle x_{D0} || 1 \rangle \underbrace{\langle D_1 || 2_{D1} \rangle}_1 - \frac{i}{\sqrt{2}} \langle x_{D0} || 1' \rangle \underbrace{\langle D_1 || 2'_{D1} \rangle}_1 - \right. \right. \\
&\quad \left. - \frac{i}{\sqrt{2}} \langle x_{D0} || 1 \rangle \underbrace{\langle D_1 || 2_{D2} \rangle}_0 + \frac{1}{\sqrt{2}} \langle x_{D0} || 1' \rangle \underbrace{\langle D_1 || 2'_{D2} \rangle}_0 - \right. \\
&\quad \left. - i \langle x_{D0} || 1' \rangle \underbrace{\langle D_1 || 2'_{D3} \rangle}_0 - i \langle x_{D0} || 1 \rangle \underbrace{\langle D_1 || 2_{D4} \rangle}_0 \right] \Big|^2 = \\
&= \frac{1}{8} \left| \langle x_{D0} || 1 \rangle - i \langle x_{D0} || 1' \rangle \right|^2 = \\
&\stackrel{(5.5j), (5.6c)}{=} \frac{1}{8} \left(\langle 1 || x_{D0} \rangle + i \langle 1' || x_{D0} \rangle \right) \left(\langle x_{D0} || 1 \rangle - i \langle x_{D0} || 1' \rangle \right) = \\
&= \frac{1}{8} \left(\underbrace{\left| \langle x_{D0} || 1 \rangle \right|^2}_{\textcircled{1}} - i \underbrace{\langle 1 || x_{D0} \rangle \langle x_{D0} || 1' \rangle}_{\textcircled{2}} + \right. \\
&\quad \left. + i \underbrace{\langle 1' || x_{D0} \rangle \langle x_{D0} || 1 \rangle}_{\textcircled{3}} + \underbrace{\left| \langle x_{D0} || 1' \rangle \right|^2}_{\textcircled{4}} \right). \tag{8.15c}
\end{aligned}$$

Here $(-i)(+i) = 1$ has been used. Finally, the probability that a photon is detected by D_0 at position x , and that about 8 ns later the partner photon is detected by D_2 , is equal to

$$\begin{aligned}
&\left| \langle x_{D0} | \langle D_2 || 8.13 \rangle \right|^2 = \\
&= \left| \frac{1}{2} \left[\frac{1}{\sqrt{2}} \langle x_{D0} || 1 \rangle \underbrace{\langle D_2 || 2_{D1} \rangle}_0 - \frac{i}{\sqrt{2}} \langle x_{D0} || 1' \rangle \underbrace{\langle D_2 || 2'_{D1} \rangle}_0 - \right. \right. \\
&\quad \left. - \frac{i}{\sqrt{2}} \langle x_{D0} || 1 \rangle \underbrace{\langle D_2 || 2_{D2} \rangle}_1 + \frac{1}{\sqrt{2}} \langle x_{D0} || 1' \rangle \underbrace{\langle D_2 || 2'_{D2} \rangle}_1 - \right. \\
&\quad \left. - i \langle x_{D0} || 1' \rangle \underbrace{\langle D_2 || 2'_{D3} \rangle}_0 - i \langle x_{D0} || 1 \rangle \underbrace{\langle D_2 || 2_{D4} \rangle}_0 \right] \Big|^2 = \\
&= \frac{1}{8} \left| -i \langle x_{D0} || 1 \rangle + \langle x_{D0} || 1' \rangle \right|^2 =
\end{aligned}$$

$$\begin{aligned}
(5.5j), (5.6c) &\stackrel{\text{---}}{=} \frac{1}{8} \left(+i \langle 1 || x_{D0} \rangle + \langle 1' || x_{D0} \rangle \right) \left(-i \langle x_{D0} || 1 \rangle + \langle x_{D0} || 1' \rangle \right) \\
&= \frac{1}{8} \left(\underbrace{\left| \langle x_{D0} || 1 \rangle \right|^2}_{\textcircled{1}} + i \underbrace{\langle 1 || x_{D0} \rangle \langle x_{D0} || 1' \rangle}_{\textcircled{2}} - \right. \\
&\quad \left. - i \underbrace{\langle 1' || x_{D0} \rangle \langle x_{D0} || 1 \rangle}_{\textcircled{3}} + \underbrace{\left| \langle x_{D0} || 1' \rangle \right|^2}_{\textcircled{4}} \right). \tag{8.15d}
\end{aligned}$$

Again $(-i)(+i) = 1$ has been used. Equations (8.15c) and (8.15d) differ from (8.10) in the different factors $1/8$ resp. $1/2$. These factors are based on the fact that the sum of the probabilities of the various possible outcomes in each experiment must equal 1. Of much greater interest are the factors $\pm i$, which now are showing up at the factors ② resp. ③, compared to (8.10). According to (A.8), a factor of plus or minus i corresponds to a phase shift of plus or minus a quarter-wavelength. (8.15c) and (8.15d) thus describe interference patterns that look as if the difference in path lengths from the upper or lower slit to the detector D_0 were a quarter-wavelength larger or smaller than the actual geometric difference.

(8.15c) is plotted as a green curve in diagram 8.10, and (8.15d) as a red curve. Due to the formal additional difference in path lengths of plus or minus $1/4$ wavelength, the interference patterns of these two curves are shifted from the center to the left resp. right.

The blue curve in diagram 8.10 is the sum of (8.15c) and (8.15d):

$$\begin{aligned}
(8.15c) + (8.15d) &= \\
&= \frac{1}{8} \left(\textcircled{1} - i \textcircled{2} + i \textcircled{3} + \textcircled{4} + \textcircled{1} + i \textcircled{2} - i \textcircled{3} + \textcircled{4} \right) = \\
&= \frac{1}{4} \left(\textcircled{1} + \textcircled{4} \right) = (8.15b) + (8.15a) \tag{8.16}
\end{aligned}$$

Due to the additional formal difference in path lengths of $\pm 1/4$ wavelength between the upper resp. lower slit and the detector D_0 , the mixed terms ② and ③ cancel each other out in this case. Only the terms ① and ④, which do not contribute to interference, remain. Thus, the blue curve is not only the sum of the red and green curves, but also the sum of the two curves (8.15a) and (8.15b), in which the origin of photon₁ and photon'₁ is precisely localized, and therefore no interference occurs.

From this beautiful experiment, much can be learned about how the positions or trajectories of quantum particles are created through measurement. I do not need to retract my repeated statements of the sort “localization is an objective fact, regardless of whether anyone takes note of the location or the localized trajectory”, but I should be more precise about when localization has taken place, and when not (yet).

- * The point of origin of the photon pair is localized as soon as either photon₂ or photon'₂ is detected by detectors D_4 or D_3 , or is irretrievably lost — for example, by hitting a wall or escaping into space.
- * As long as photon₂ or photon'₂ has neither been detected nor irretrievably lost, the point of origin of the photon pair is *not yet* localized. While the existence of photon₂ resp. photon'₂ offers the possibility of localizing the point of origin, this possibility has not yet been realized.
- * If we succeed in capturing and detecting photon₂ or photon'₂ using a suitable interferometer (in our example, with detectors D_1 or D_2), this definitively rules out the possibility of localization. In this case, the point of origin of the two photons was *never* localized, i. e. there is no localization that could be “erased”.

Once the localization has been established, even the quantum

eraser cannot erase it. At most, one could say that the quantum eraser erases the *possibility of localization*. In this respect, the name “quantum eraser” is misleading and somewhat poorly chosen, but out of respect for the discoverers of the effect (Marlan Scully and Kai Drühl [61]), physicists have retained this name.

At this point, we should recall Bohr’s admonition (formulated above as assumption C) regarding the holistic approach to quantum phenomena. The *individual quantum phenomenon* of the two SPDC-generated photons is not yet completed as long as both photons have not either been detected by (classically described!) measuring devices or have been definitively lost. As long as this has not yet happened, it does not make sense to speak of the parts of the quantum phenomenon, nor to say that one photon is localized by the other.

The phrasing “...both photons are either ...detected ...or have been definitively lost” is interesting. This suggests that the definitive loss of a photon is, in a certain sense, equivalent to its detection by a measuring device. Detection and definitive loss are indeed equivalent in that neither can be reversed. This aspect, discussed by physicists under the term *decoherence*, is so important for the understanding of quantum phenomena that I will return to it later and devote the entire section 9.4 to it.

8.4 The interaction-free Light-Barrier

Fig. 8.11 on the following page shows a simple light barrier. In 8.11(a), the light path is free, thus the detector D_F detects the light. If D_F does not detect any light, then one can conclude that — as sketched in 8.11(b) — there must be an object between the turquoise-colored light source and the detector that is blocking the light path. Can a light barrier also be operated with a single photon? Fig. 8.12 illustrates how a 1-photon light barrier could be



Fig. 8.11 : A conventional light barrier

constructed. Using the SPDC process (explained in Section 3.5), photon pairs are generated, which we will call photon_1 and photon_2 as usual. When the detector D_G detects photon_1 , we wait 10 ns to see whether photon_2 is detected by D_F ; then the experiment ends.

If D_F detects photon_2 within the 10 ns window after D_G has triggered, we know that the light path is clear. The reverse is not true: If D_F does not respond within the 10 ns window, this could be either because an object is blocking the light path, or because D_F did not respond even though photon_2 reached the detector. This cannot be ruled out because the efficiency of photon detectors is well below 100%.

The 1-photon light barrier therefore is not a particularly impressive device. It produces a remarkable effect, however, when configured in a slightly more complex way, namely as arm A of the two arms A and B of an interferometer. This is illustrated in fig. 8.13.

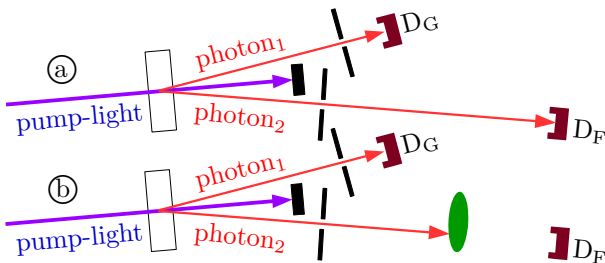


Fig. 8.12 : A 1-photon light barrier

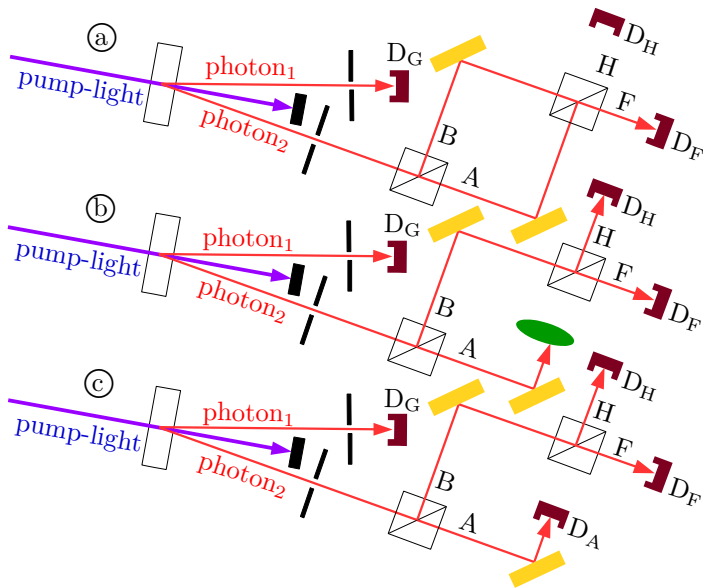


Fig. 8.13: 1-photon light barrier within an interferometer

First the interferometer is carefully adjusted, so that photon_2 always reaches detector D_F , if the light path is unobstructed — see fig. 8.13(a) —, while detector D_H never triggers when the light path is unobstructed. In section 2.1, it was explained that this is the case, for example, when the lengths of light paths A and B are exactly identical. Then the partial wave reflected from path A into path F constructively interferes with the partial wave transmitted from path B into path F. And the partial wave transmitted from path A into path H destructively interferes with the partial wave reflected from path B into path H.

Once the interferometer has been adjusted, the experiment begins: We wait until D_G responds for the first time. Then we record whether D_F or D_H responds within 10 ns, or whether neither de-

detector responds. This concludes the experiment.

If the light path is clear, as shown in 8.13(a), then — because the efficiency of the detectors is well below 100% — D_F will sometimes trigger in a series of experimental runs, sometimes no detector will trigger, but D_H will never trigger.

If, on the other hand, an obstacle is present in path A as shown in 8.13(b), then also detector D_H will sometimes be triggered, because the partial wave transmitted from path A, which previously had caused destructive interference in path H, now is absent.

If D_F triggers, or if none of the detectors triggers, then the result is ambiguous: perhaps an obstacle is blocking the light barrier (i. e. path A), or perhaps not. The result is unambiguous, however, if D_H triggers: this result can only occur if path A is blocked.

A widespread interpretation (which I will, however, criticize immediately) of experimental runs, in which D_H detects the photon, is this:

- I** The activation of D_H proves that the light barrier is blocked by an obstacle. Therefore, the photon can impossibly have reached the detector via path A; it must have taken path B. But how, then, could the 1-photon light barrier have detected the obstacle, even though the only photon₂ involved here didn't even come close to the obstacle?

If one interprets the process in this way, it stands to reason to call the light barrier “interaction-free”, since no interaction seems to have taken place between the photon and the obstacle.

The concept of this interaction-free light barrier was published in 1993 by Avshalom Elitzur and Lev Vaidman [62, 63]. In 1995, Kwiat et al. [64] verified in the laboratory, that the method actually works. In doing so, the experimenters replaced the obstacle sketched in green in fig. 8.13(b) by an additional photon detector D_A , as displayed in fig. 8.13(c). The advantage is that the processes can

be observed and verified even more precisely with this additional detector.⁷² Therefore, I will discuss in the following the light barrier with this additional detector.

The interpretation **I** is not without alternative, and it does not align with the ideas developed in the previous chapters regarding the interpretation of the 1-particle-interference. Let us revisit in this context the “time-sketches” of fig. 3.10 on page 69.

The split position of the photon moves toward the detectors D_R and D_T . Only in the moment (time ④) when the photon is detected by D_R does its position suddenly collapse to the active area of D_R . In the same way, we can illustrate step-by-step how the position of photon_2 moves through the “interaction-free light barrier”. This is shown in fig. 8.14 on the following page.³⁰

Let us formulate this perspective — which is an alternative to interpretation **I** — as interpretation **J**:

J At the beam splitters, the path of photon_2 is split into both possible paths. Therefore, part of its path temporarily touches the obstacle in path A (i. e. the detector D_A). Only if the detector D_A does not respond (and also does not absorb the photon without responding) even though it was touched by photon_2 , the path of the photon is confined to path B. And only when D_H triggers (not earlier!) does the location of photon_2 suddenly shrink to the active area of D_H .

Both interpretations are consistent with the observed result, namely that the light barrier recorded the obstacle by detecting the photon with detector D_H . In this respect, neither interpretation is “wrong”.

The interpretation **J** has the advantage of sparing us an unsolvable puzzle, namely the question of why the light barrier was able

⁷² Elitzur and Vaidman [62] discussed the Mach-Zehnder-interferometer sketched in fig. 8.13, whereas in the experiment by Kwiat et al. [64] a Michelson-interferometer was used. This difference, however, is irrelevant to our investigation.

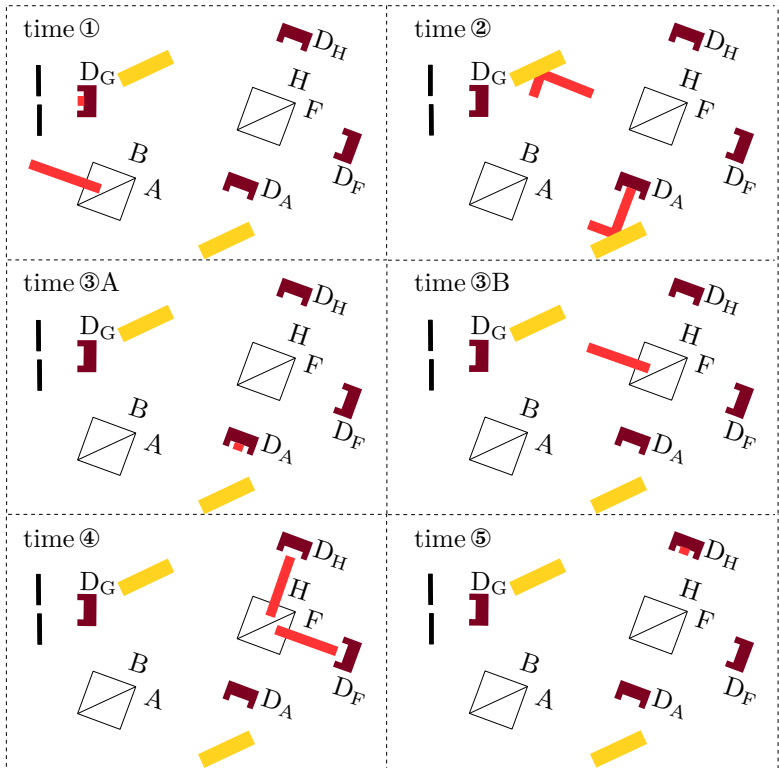


Fig. 8.14: The position of photon₂

time ①: The red sketched position of photon₁ is created in the moment in which it is detected by D_G .

time ②: The position of photon₂ is split onto the arms A and B of the interferometer. It already touches D_A , but the detector has not (yet) triggered.

time ③A: D_A has registered the photon. Thereby its position has collapsed to the active surface of D_A . The experiment is finished.

time ③B: D_A die *not* register the photon. Thereby its position collapsed to path B.

time ④: The photon's position is split to paths F and H.

time ⑤: D_H has registered the photon. Thereby its position has collapsed to the active surface of D_H .

to detect the obstacle even though the only photon was traveling along path B. One could argue that an advantage of **I** is that this interpretation spares us the splitting of the photon's location across different paths. It does so, however, only if there is an obstacle in the light barrier. If the light barrier is clear, we still need the splitting of the photon's location in order to interpret the single-particle interference. Therefore, to me interpretation **J** seems to be much more elegant and plausible overall.

Let's try to underpin what is suggested by the sketches in fig. 8.14 with sound quantum-theoretical considerations: At the first beam splitter, photon₂ is transmitted or reflected with equal probability; that is, its state vector becomes

$$|\text{photon}_2\rangle = \sqrt{\frac{1}{2}} \left(|A\rangle - i|B\rangle \right). \quad (8.17a)$$

$|A\rangle$ is the state vector of a photon that takes path A. $|B\rangle$ is the state vector of a photon that takes path B. The factor $-i = -\sqrt{-1}$ arises because at a beam splitter there is always a phase shift of a quarter wavelength between the transmitted and reflected particles. According to (A.8), this phase shift is represented in theory by the factor $-i$. The factor $\sqrt{1/2}$ is necessary so that the projection amplitude of this entangled state vector onto itself is equal to 1, as it must be.

On path A, the photon either strikes D_A (if this detector is installed) or continues on its path undisturbed (if D_A is not installed). This can be written as follows:

$$\begin{aligned} |A\rangle &= c_1 |A, D_A\rangle + c_2 |A\rangle \\ &\text{with } |c_1|^2 + |c_2|^2 = 1 \end{aligned} \quad (8.17b)$$

$|A, D_A\rangle$ is the state vector of the photon that strikes detector D_A via path A. If the detector blocks path A, then $c_1 = 1$ and $c_2 = 0$. If path A is free, then $c_1 = 0$ and $c_2 = 1$. The more general

notation (8.17b) saves us from having to write separate equations for each case.

If the photon reaches the second beam splitter, it will be transmitted or reflected with equal probability:

$$|A\rangle = \sqrt{\frac{1}{2}} \left(|H_A\rangle - i|F_A\rangle \right) \quad (8.17c)$$

$$|B\rangle = \sqrt{\frac{1}{2}} \left(|F_B\rangle - i|H_B\rangle \right) \quad (8.17d)$$

The factor $-i$ was explained below (8.17a). $|F_A\rangle$ is the state vector of a photon that was reflected into path F after coming from path A. $|F_B\rangle$ is the state vector of a photon that was transmitted into path F after coming from path B. $|H_A\rangle$ is the state vector of a photon that was transmitted from path A to path H. $|H_B\rangle$ is the state vector of a photon that was reflected from path B to path H.

Now we combine the equations (8.17):

$$\begin{aligned} |\text{photon}_2\rangle &= \sqrt{\frac{1}{2}} \left[c_1 |A, D_A\rangle + \right. \\ &\quad \left. + c_2 \sqrt{\frac{1}{2}} \left(|H_A\rangle - i|F_A\rangle \right) - i \sqrt{\frac{1}{2}} \left(|F_B\rangle - i|H_B\rangle \right) \right] = \\ &= \sqrt{\frac{1}{2}} c_1 |A, D_A\rangle + \frac{1}{2} \left(c_2 |H_A\rangle - i c_2 |F_A\rangle - i |F_B\rangle - |H_B\rangle \right) \quad (8.18) \end{aligned}$$

In the last summand, $(-i)^2 \stackrel{(5.5e)}{=} -1$ has been used.

The efficiency of photon detectors is well below 100%, i. e. not every photon that reaches a detector is actually detected. For simplicity, we assume for the moment being that the three photon detectors do have an efficiency of 100%. Later we then will have to correct our results accordingly. Assuming 100% efficiency, the probability $P(D_H)$ that the detector D_H registers the photon is, according to Born's rule (5.19b), equal to the square of the projection amplitude of (8.18) onto the eigenvector $|D_H\rangle$ of this detector:

$$\begin{aligned}
 P(D_H) = & \left| \langle D_H || \text{8.18} \rangle \right|^2 = \frac{1}{4} \left| \sqrt{2} c_1 \underbrace{\langle D_H || A, D_A \rangle}_0 + c_2 \langle D_H || H_A \rangle - \right. \\
 & \left. - i c_2 \underbrace{\langle D_H || F_A \rangle}_0 - i \underbrace{\langle D_H || F_B \rangle}_0 - \langle D_H || H_B \rangle \right|^2 \quad (8.19a)
 \end{aligned}$$

The photon cannot possibly reach detector D_H if it was intercepted by detector D_A or if it took path F at the second beam splitter. Therefore, three projection amplitudes are zero. In the same way, we find the probability $P(D_F)$ that detector D_F registers the photon, and the probability $P(D_A)$ that detector D_A (i. e. the obstacle in the light barrier) registers the photon:

$$\begin{aligned}
 P(D_F) = & \frac{1}{4} \left| \sqrt{2} c_1 \underbrace{\langle D_F || A, D_A \rangle}_0 + c_2 \underbrace{\langle D_F || H_A \rangle}_0 - \right. \\
 & \left. - i c_2 \langle D_F || F_A \rangle - i \langle D_F || F_B \rangle - \underbrace{\langle D_F || H_B \rangle}_0 \right|^2 \quad (8.19b)
 \end{aligned}$$

$$\begin{aligned}
 P(D_A) = & \frac{1}{4} \left| \sqrt{2} c_1 \langle D_A || A, D_A \rangle + c_2 \underbrace{\langle D_A || H_A \rangle}_0 - \right. \\
 & \left. - i c_2 \underbrace{\langle D_A || F_A \rangle}_0 - i \underbrace{\langle D_A || F_B \rangle}_0 - \underbrace{\langle D_A || H_B \rangle}_0 \right|^2 \quad (8.19c)
 \end{aligned}$$

If an obstacle blocks path A, then $c_1 = 1$ and $c_2 = 0$. If no obstacle blocks path A, then $c_1 = 0$ and $c_2 = 1$. Therefore,

if path A is free:

$$P(D_H) \stackrel{(8.19a)}{=} \frac{1}{4} \left| \langle D_H || H_A \rangle - \underbrace{\langle D_H || H_B \rangle}_{= \langle D_H || H_A \rangle} \right|^2 = 0 \quad (8.20a)$$

$$\begin{aligned}
 P(D_F) &\stackrel{(8.19b)}{=} \frac{1}{4} \left| -i \langle D_F || F_A \rangle - i \underbrace{\langle D_F || F_B \rangle}_{=\langle D_F || F_A \rangle} \right|^2 = \\
 &= \frac{1}{4} \underbrace{\left| -i \right|^2}_1 \underbrace{\left| 2 \right|^2}_4 \underbrace{\left| \langle D_F || F_A \rangle \right|^2}_1 = 1
 \end{aligned} \tag{8.20b}$$

$$P(D_A) \stackrel{(8.19c)}{=} 0 \tag{8.20c}$$

if detector D_A blocks path A:

$$P(D_H) \stackrel{(8.19a)}{=} \frac{1}{4} \underbrace{\left| -1 \right|^2}_1 \underbrace{\left| \langle D_H || H_B \rangle \right|^2}_1 = \frac{1}{4} \tag{8.20d}$$

$$P(D_F) \stackrel{(8.19b)}{=} \frac{1}{4} \underbrace{\left| -i \right|^2}_1 \underbrace{\left| \langle D_F || F_B \rangle \right|^2}_1 = \frac{1}{4} \tag{8.20e}$$

$$P(D_A) \stackrel{(8.19c)}{=} \frac{1}{4} \underbrace{\left| \sqrt{2} \right|^2}_2 \underbrace{\left| \langle D_A || A, D_A \rangle \right|^2}_1 = \frac{1}{2} \tag{8.20f}$$

In state $|H_B\rangle$ the photon was reflected by both beam splitters, whereas in state $|H_A\rangle$ it was transmitted by both beam splitters. Therefore, there is a phase shift of half a wavelength between these two state vectors (destructive interference), which is the cause of the factor -1 in (8.20a); see (A.8).

In state $|F_B\rangle$ the photon was reflected at the first beam splitter and transmitted at the second beam splitter, whereas in state $|F_A\rangle$ it was transmitted at the first beam splitter and reflected at the second beam splitter. Therefore, these two state vectors are in phase (constructive interference); both have the phase factor $-i$ in (8.20b) due to the single reflection, compare (A.8).

The quantum-theoretical description can be viewed as an argument in favor of interpretation **J**: While in case $c_1 = 1, c_2 = 0$ (i. e. if path A is blocked by an obstacle), in

$$P(D_H) \stackrel{(8.20d)}{=} \frac{1}{4} \underbrace{|-1|^2}_1 \underbrace{|\langle D_H || H_B \rangle|^2}_1 = \frac{1}{4}$$

only the state vector $|H_B\rangle$ of a photon shows up, which arrived at detector D_H via path B, this term becomes only relevant once the photon has arrived at this detector. Until that moment, the state vector of photon₂ is

$$|\text{photon}_2\rangle \stackrel{(8.18)}{=} \sqrt{\frac{1}{2}} |A, D_A\rangle + \frac{1}{2} \left(-i|F_B\rangle - |H_B\rangle \right).$$

This state vector also includes the state vector $|A, D_A\rangle$ of the photon that reaches the obstacle in the light barrier (i. e. the detector D_A). This state vector suggests the interpretation that the trajectory of photon₂ is not yet restricted to path B, but rather extends across both arms of the interferometer, as sketched in fig. 8.14.

The name “interaction-free light barrier” could be justified within the framework of interpretation J by the fact that the obstacle in the light barrier (the detector D_A) is clearly detected in one-quarter of the experimental runs by the response of D_H , without D_A responding (even if the detectors have an efficiency of 100 %!). Elitzur and Vaidman have emphasized in their article [62] that they do not understand the term “interaction-free” in the sense of “ D_A is not touched by the photon”, but rather in the sense of “the detector D_A does not respond, even though it is detected as an obstacle”.

But why is D_A — even at 100 % detector efficiency — according to (8.20f) only triggered in half of the experimental runs, even though the split position of the photon touches this detector in *every* experiment run according to interpretation J? This is because only half of the photon’s position touches this detector; the other half is on path B. Even with 100 % detector efficiency, half a position can trigger a detector in only half of the experimental runs.

With 100% detector-efficiency, the probabilities add to 1:

if detector D_A blocks the light barrier:

$$P(D_H) \stackrel{(8.20d)}{=} \frac{1}{4} \quad , \quad P(D_F) \stackrel{(8.20e)}{=} \frac{1}{4} \quad , \quad P(D_A) \stackrel{(8.20f)}{=} \frac{1}{2}$$
$$P(D_H) + P(D_F) + P(D_A) = 1$$

Thus never two — or even all three — detectors must trigger in the same experimental run. If more than one detector could trigger in the same experimental run, then the sum of the three probabilities would be greater than 1. Nature can only achieve this correlation or anticorrelation (with 100% detector-efficiency, at least one detector must trigger in every experimental run, but never may two or three detectors trigger) only by coordinating the three detectors non-locally. We have already observed that nature indeed is acting non-locally in numerous other quantum phenomena.

9 Interpretations

Due to ingenious guessing, Heisenberg and Schrödinger had discovered in 1925/26 the mathematical formalism that correctly describes quantum phenomena. But few physicists were (and are) entirely satisfied with an abstract formalism. Most desire more tangible, intuitive explanations for these phenomena. Such more intuitive explanations are referred to as *interpretations* of quantum phenomena.

9.1 The Copenhagen Interpretation

We know that electrons are no particles. This is proved by every interference experiment with electrons. We also know that electrons are no waves. If the double-slit interference experiment is modified by placing detectors immediately behind each slit, then only one of the two detectors ever registers the complete electron. There never arrive only parts of an electron at both detectors, which would have to happen if electrons were waves.

It is easy to say what electrons are *not*. Apparently it is much more difficult to state positively what electrons and other quantum objects actually are. In winter 1926/27, Heisenberg and Bohr came to the conclusion that it is not only difficult, but simply impossible, to state positively what quantum objects are.

Five years earlier, while walking in summer 1922 across the Hainberg near Göttingen, Heisenberg had asked: “If we don’t actually

have a language with which to talk about these structures, will we then ever understand atoms at all?” And Bohr had replied: “Yes, but in doing so, we will first learn what the word ‘to understand’ does mean.” This took the two physicists now seriously. They considered what the word “to understand” could reasonably mean in the analysis of quantum phenomena. The result of their deliberations, known as the “Copenhagen Interpretation”, looks something like this:

When we say that we have understood something, we mean that we have succeeded in building a bridge between the objective reality and our human capacity for understanding. The physical theory is the formal expression of our understanding. On the one hand, the theory reflects the objective reality; on the other hand, it reflects the nature of human thought. Our human capacity for thought, in turn, did not suddenly fall from the sky, but has developed over the many, many millions of years of our evolution. And during this time, it has proven itself in interaction with the environment; otherwise, we would not be here today. Therefore, we can trust that our way of thinking, our cognitive abilities, and the concepts we have developed over this long period of time at least adequately represent those phenomena in our everyday environment that were important during evolution. The concepts of classical physics are the refined and precise form of our everyday concepts, but they are not entirely different ones. Only these fit into our brains; we have no others, we will not get any others, and we must see how we can make the best of our capabilities for thinking and understanding.

When the concepts of classical physics are applied to quantum objects, they fail. The self-interference of electrons, for example, cannot be forced into the framework of classical physics. Because Bohr and Heisenberg were convinced that concepts other than the classical ones do not fit into human brains, they should have given up at this point — had they not come up with a trick.

While the concepts of classical physics do not fit to quantum objects, they fit perfectly to a large part of the world. In particular, they fit to the measuring instruments (detectors, slits, mirrors, apertures, etc.) used in the observation of quantum objects. Although these devices are all composed of atoms and could therefore be described as giant molecules using the methods of quantum theory, they must — this is the crucial trick in the Copenhagen Interpretation — be described using the methods and concepts of classical physics. Only in this way, according to Heisenberg and Bohr, do we gain the tools with which we can attempt an analysis of quantum phenomena. If we were to regard the measuring instruments as quantum objects as well, we would saw off the branch on which we sit. Then there would be no starting point left for our human cognitive faculties; we would be completely helpless.

In a lecture [65] (which, by the way, I highly recommend to all readers of this book who wish to learn about the Copenhagen Interpretation firsthand), Heisenberg explained in 1955:

“The concepts of classical physics are merely a refinement of the concepts of everyday life, and form an essential part of the language that serves as the foundation for all natural sciences. Our actual situation in the natural sciences is such that we do in fact use — and must use — classical concepts to describe our experiments; for otherwise we could not communicate with each other. And the task of quantum theory was precisely to interpret the experiments theoretically on this basis. It makes no sense to discuss what could be done if we were beings other than what we actually are. At this point, we must realize that, as v. Weizsäcker put it, ‘nature is older than man, but man is older than natural science.’ The first part of the sentence justifies classical physics with its ideal of complete objectivity. The second part explains why we cannot escape the paradox of quantum

theory; namely, why we cannot escape the necessity of using classical concepts.”

Quantum objects only become phenomena in interaction with measuring instruments. In some cases, the classical measuring instrument may be as seemingly simple as the human eye.⁷³ The entire arrangement of devices — such as mirrors, slits, detectors, and so on — is an inseparable part of the quantum phenomenon, just as the atoms, photons, and other quantum objects are that move within this arrangement, and whose existence only (!) comes to the experimenter’s attention through this arrangement. A physicist never deals with an isolated quantum object; quantum objects are always and with no exception embedded in a classically describable framework, and only thereby become quantum phenomena. Without this framework, they would simply not be phenomena⁷⁴, and physics would have no reason to concern itself with a quantum object that is merely imagined but not observed.

If the apparatus is set up such that an electron interferes with itself, then, according to the Copenhagen Interpretation, it may be conceived of as a wave, and one may use all the concepts that classical physics has developed for waves. If it is observed with a particle detector, then it may be conceived of as a particle, and one may use all the concepts that classical physics has developed for particles. Heisenberg’s indeterminacy relation (8.1) expresses the fact that no contradictions can ever arise in this context. For it is impossible to construct an apparatus that can simultaneously measure a quantum object as a particle with a precisely defined position ($\Delta x \approx 0$) and as a wave with a precisely defined wavelength ($\Delta p \stackrel{(4.2b)}{=} \Delta(h/\lambda) \approx 0$). Sequentially that is possible: In the experiment by Thorn et al., the photon is first created as a wave that

⁷³ In 2016, it was demonstrated for the first time that a well-adapted human eye can indeed perceive individual photons. [66]

⁷⁴ The Greek word phenomenon means “the visible”.

splits at the beam splitter. Then it is converted by the detector into a particle with a precisely defined position, see fig. 3.10 on page 69. Sequentially that is possible, but not simultaneously.

The absurdity depicted in fig. 9.1 is intended to illustrate this idea. Of course, we can define the term “beipe” for this absurdity. Then it has a name, but what does that help? The name does not make it any less absurd. If the drawing is covered so that only the right part is visible, however, then the beipe takes the form of three sensible pipes. If only the left part is visible, then the beipe takes the form of two sensible beams.



Fig. 9.1: Two beams? Three pipes? A beipe!

For the absurd quantum objects we also have names, such as “electron”, “photon”, or “atom”. These names are just as unhelpful as the name “beipe”. In order to associate “reasonable” concepts with electrons, photons, or atoms, we must observe them in an appropriate way of viewing. This “appropriate way of viewing” is automatically ensured by the measuring instruments, which impose either the properties of waves or the properties of particles on these objects. According to Heisenberg’s undeterminacy principle (8.1), the measuring instruments can either create a precisely defined location for the quantum object ($\Delta x \approx 0$); then it has the properties of a particle. Or they can create a precisely defined momentum — and thus simultaneously a precisely defined wavelength ($\Delta p \stackrel{(4.2b)}{=} \Delta(h/\lambda) \approx 0$) — of the quantum object; then it has the properties of a wave. But not both at the same time.

Bohr therefore emphasized throughout his life that concepts such as wave and particle are not contradictory but *complementary* when describing quantum phenomena. Only by combining complementary images and concepts, so his *credo*, can humans expand their evolutionarily limited cognitive capabilities to the point where they can, after all, fully comprehend quantum phenomena.

If one sets aside the classical apparatus and reflects on the nature of an electron, seeking to understand what the electron “in itself” is, one is faced with an enigmatic mystery. According to Heisenberg and Bohr, however, this is a non-physical question, because an electron that is not embedded in a classical apparatus cannot be an object of physical observation. Only an isolated quantum object without a classical framework would seem incomprehensible to us. But all observed quantum phenomena can be understood precisely because of their observability and within the classical framework that brings about their observability.

According to the Copenhagen Interpretation, quantum phenomena must always be considered holistically (as *individual* phenomena), including all apparatus and measuring instruments. If one asks questions such as “is an electron a wave or a particle?” or “is a photon split by a beamsplitter, or does it choose one direction or the other?” without referring to the entire experimental setup and incorporating it into the description, one quickly ends up in a quagmire of confusing absurdities. But as soon as one acknowledges that we can speak meaningfully about a quantum phenomenon only as a whole, all apparent contradictions are instantly resolved.

The classically described framework in which atoms, electrons, and photons are embedded is an integral, inseparable component of every quantum phenomenon; in Aristotelian terms, it is the “form” of the phenomenon.⁷⁵ This is a fortunate coincidence: Everything

⁷⁵ An essay (in German) on the influence of Aristotle’s philosophy on the Copenhagen Interpretation of quantum phenomena can be found here: [67]

we can observe, we can also understand. And everything we can understand, we can also observe. In both directions, it is the classical framework of quantum phenomena — that is, the classically described measuring instruments — that brings about the observability and comprehensibility of quantum phenomena.

A phenomenon without form is just as unthinkable an absurdity as a one-sided coin. Such a thing does not fit into a human brain, and according to Bohr and Heisenberg it is not a reasonable goal to strive for an understanding of one-sided coins or of atoms, electrons, or photons outside of a classical framework. Individual quantum phenomena are the smallest units we can meaningfully study. According to the Copenhagen Interpretation, any further conceptual division of our objects of study does not lead to further insight, but only to useless confusion.

From my description of the Copenhagen Interpretation so far, it appears as though the classical description of measuring instruments is merely a stopgap, a crutch on which humans depend due to their evolutionarily limited cognitive capabilities. But in his 1955 lecture, cited above [65], Heisenberg also presents a positive argument why for measuring instruments the classical description is in fact more appropriate than a quantum-theoretical description. I will comment on that argument, which is discussed under the heading “decoherence” in the physics literature, in section 9.4.

In the same lecture [65], Heisenberg also addresses the question of whether — or to what extent — quantum theory still does meet the ideal of objectivity in science. For in the investigation of quantum phenomena, one never observes nature “in itself”, but rather nature as it is subjected — through the experimenter’s arbitrary choice of the type and design of the measuring instruments — to our respective research questions. The measuring instruments in general do not simply determine what was already the case, but rather *create* the reality that we determine with their help.

Heisenberg concludes that quantum theory does meet the ideal of objectivity “as far as possible”, which in plain language means: Not much remains of this ideal. It is simply impossible to observe nature without observing it. Bohr liked to express this in the flowery phrase that “in the drama of life, we always are both spectators and participants at the same time.”

One can understand the unique significance that measurements hold for quantum phenomena by considering the roles that musicians and listeners play in the phenomenon of music. The listeners correspond to the classical physics conception of measuring instruments. They record everything precisely without noticeably affecting the object being measured. If one regards music as an analogy for quantum phenomena, however, then one must designate the musicians — but not the listeners — as “measuring instruments”. For measuring instruments exert a distinct and non-negligible effect that shapes the manifestation of the quantum phenomenon, just as the musicians shape the manifestation of the music.

When the musicians of a string quartet set aside their bows for a pizzicato, and pluck the strings with their fingers, then the music suddenly sounds completely different. But no one, of course, demands that we should finally determine objectively what “true” or “actual” music “in itself” sounds like when it is characterized neither by the use of violin bows nor by plucking with the fingers. For music arises in the first place through the interaction of musical instruments, bows, and the musicians’ fingers, and it would be completely meaningless to speak of music without these prerequisites. According to the Copenhagen Interpretation of quantum phenomena, it would be just as pointless to ask what properties electrons or other quantum objects have “objectively” and “in themselves”, if the properties of waves or particles are not imposed on them through the use of appropriately constructed measuring

instruments.

Interestingly, even in classical physics there are phenomena that are only brought about by observation and do not exist without it. A well-known example is the rainbow. The objective fact is that sunlight shines into a rain shower, is refracted — depending on the angle of incidence — by each individual drop, broken down into its various colors, and partially reflected. All colors are scattered more or less evenly in all directions. A “bow” does not appear at all in an objective, complete description of the facts.

A bow is showing up only if an observer, from his vantage point, selectively perceives only a tiny fraction of the scattered light. And the observer’s vantage point also determines the points on the Earth’s surface where the base of the rainbow is located. Another observer, a few hundred meters away, also sees a rainbow. But one cannot really say that it is the same rainbow, and in any case, the two observers see the bases of their respective arcs at different points on the Earth’s surface. If no observer is there to look, then the sun and the raindrops can try as hard as they want, they will not produce a rainbow.⁷⁶

Nevertheless, the rainbow is a purely classical phenomenon because it is possible to analyze in minute detail what the objective (independent of observation) facts are — namely, the refraction of sunlight and its more or less uniform scattering in all directions, and which characteristics are produced by the act of observation — namely, the shape of the arc and the position of its base points.

⁷⁶ On page 199 I quoted from Pais’s account: “Einstein suddenly stopped, turned to me and asked whether I really believed that the moon exists only when I look at it.” [55, page 907] If Pais had replied at that moment: “But remember the rainbow. This phenomenon only exists if somebody is looking at it.”, then Einstein might have acknowledged after all that some properties of observed objects — such as the arched shape of a rainbow or the location of an atom — are not entirely independent of the nature of the observation, and consequently exist not entirely independent of the fact that they are observed.

In contrast, it is characteristic of quantum phenomena that a clear distinction between objective facts and the influence of the observer resp. the measuring instruments is never possible. This is precisely why the state vector always contains an inseparable mixture of both. There is an ongoing debate among physicists as to whether a reasonable theory should not at least allow for a clear separation between “objective” and “subjective” (observation-induced) aspects of the phenomena. According to Heisenberg and Bohr, however, the inseparable mixing of objective and subjective components in the state vector arises from the very nature of the issue and is absolutely unavoidable.

9.2 John v. Neumann’s Collaps

Methods are called heuristic when they ultimately lead to the correct solution despite following dubious or even flawed paths. The heuristic tricks of brilliant theoretical physicists have always aroused displeasure among serious mathematicians. That was no different during the development of the nascent quantum theory. What Heisenberg, Jordan, Dirac, Schrödinger, and Pauli had conjured up within a few years literally cried out for consolidation by a mathematical expert.

John⁷⁷ v. Neumann (1903–1957), a mathematician not less brilliantly gifted⁷⁸ than the physicists mentioned above, devoted himself to this task. In 1932, he published his book “Mathematical

⁷⁷ In his youth in Budapest, he was called János. During his studies in Switzerland and Germany, he went by Johann. He adopted the first name John in the United States, where he became a leading pioneer of early computer technology.

⁷⁸ Reportedly, the six-year-old János astonished his family by being able to divide eight-digit numbers in his head at high speed.

Foundations of Quantum Mechanics” [68], which remained the globally recognized standard work in this field for many years.

Von Neumann worked out the mathematical structure of the theory, but did not need to make many corrections to the physicists’ results. Despite occasionally dubious methods⁷⁹ the physicists had essentially guessed everything correctly. Only on one point did v. Neumann break completely new ground: that was on the subject of measurement.

In their Copenhagen Interpretation, Heisenberg and Bohr had insisted that measuring instruments must be described by the methods of classical physics. This seemed to make no sense at all to v. Neumann. After all, measuring instruments are also made up of atoms, so quantum theory should consequently describe them as giant molecules, at least in principle.

In equation (5.17) the state vector of a neutron is specified, which, after passing through a double slit, strikes a detector that localizes the neutron at one of 87 possible positions:

$$|\text{neutron}\rangle \stackrel{(5.17)}{=} \sum_{j=1}^{87} |x_j\rangle \underbrace{\left(l \langle x_j || S_{2l} \rangle + r \langle x_j || S_{2r} \rangle \right)}_{c_j} \quad (9.1)$$

The notation c_j has been introduced here for the projection amplitudes. The summation symbol \sum has been explained below equation (5.17). The vectors $|x_j\rangle$ are the 87 eigenvectors of the detector. If, for example, the detector responds at position x_{31} , then it prepares the neutron in the state $|x_{31}\rangle$, meaning that it

⁷⁹ for physicists: Dirac’s delta function, in particular, deeply troubled the mathematicians. Only two decades later they succeeded in providing a mathematical justification for this “improper” function within the framework of the theory of distributions. J. v. Neumann mentions the delta-function only once in his book, namely in the introduction. There he notes that quantum theory can also be formulated without this strange construct, and he then consistently does so.

localizes it at position x_{31} . According to Born's rule (5.19b), the probability $W(x_{31})$ for localization at this point is equal to the square of the projection amplitude of the state vector $|\text{neutron}\rangle$ onto the eigenvector $|x_{31}\rangle$ of the measuring device:

$$W(x_{31}) \stackrel{(5.18)}{=} |c_{31}|^2 = |l \langle x_{31} || S2_l \rangle + r \langle x_{31} || S2_r \rangle|^2 \quad (9.2)$$

The state vectors $|x_{31}\rangle$ — resp. $|x_j\rangle$ in general — do not describe the measuring device; rather, they describe the neutron that is prepared by the measuring device in one of these states. We have not defined any state vector for the measuring device at all so far, since, according to the Copenhagen Interpretation, it must be described by the methods of classical physics.

At this point, v. Neumann took a different approach. He assigned the state vector $|M_j\rangle$ to the measuring device when it registers a reading at position x_j . Before the measurement, while the device has not yet registered anything at all, it is in the state $|M_0\rangle$. According to v. Neumann's conception, the measurement then proceeds in two steps. In the first step, the state vectors of the measuring device and the object (i.e., the neutron in our example) become entangled with each other:

$$\left. \begin{array}{l} |\text{neutron}\rangle \stackrel{(9.1)}{=} \sum_{j=1}^{87} c_j |x_j\rangle \\ |\text{meas. dev.}\rangle = |M_0\rangle \end{array} \right\} \xrightarrow{\text{step 1}} \\ \xrightarrow{\text{step 1}} |\text{neutron \& meas. dev.}\rangle = \sum_{j=1}^{87} c_j |x_j\rangle |M_j\rangle \quad (9.3a)$$

In general, several or all of the amplitudes c_j are nonzero. But no one has ever observed a measuring device that displays different values M_j simultaneously; every (properly functioning) measuring device displays a uniquely determined value M_k . (In the example

just mentioned, $k=31$.) Thus there must be a second step in the measurement process, in which the entangled state vector (9.3a) “collapses” to a specific value:

$$\begin{aligned} |\text{neutron \& meas. dev.}\rangle &= \sum_{j=1}^{87} c_j |x_j\rangle |M_j\rangle \xrightarrow{\text{step 2}} \\ &\xrightarrow{\text{step 2}} |x_k\rangle |M_k\rangle \end{aligned} \quad (9.3b)$$

Often step 2 is formulated in terms of the projection amplitudes:

$$c_j \xrightarrow{\text{step 2}} \begin{cases} c_j = 1 & \text{if } j = k \\ c_j = 0 & \text{if } j \neq k \end{cases}$$

The probability that the entangled state vector (9.3a) collapses precisely onto the vector $|x_k\rangle |M_k\rangle$, but not onto any other vector, is according to Born’s rule

$$W(k) \stackrel{(5.18)}{=} |c_k|^2 .$$

In example (9.2), k was 31.

But why does the collapse occur in the first place? What is fundamentally different about measuring devices compared to other molecules, in which no collapse occurs? v. Neumann’s answer: There is absolutely nothing different about measuring devices compared to other molecules. (9.3b) is merely a shorthand notation for a much more complicated process, which actually proceeds as follows. After the first step (9.3a), in which neutron and measuring device got entangled, the experimenter looks at the device’s display. This creates an image of the display on the retina of his eye. Since the retina is made up of atoms, it is also described using quantum theory. We call it’s eigenvectors $|R_j\rangle$. Before the experimenter looks at the display, the retina of his eyes is in the state $|R_0\rangle$.

Because the image of the display appears on the retina, the display and the retina get entangled:

$$\begin{aligned}
 & \left. \begin{aligned}
 |\text{neutron \& meas. dev.}\rangle & \stackrel{(9.3a)}{=} \sum_{j=1}^{87} c_j |x_j\rangle |M_j\rangle \\
 |\text{retina}\rangle & = |R_0\rangle
 \end{aligned} \right\} \xrightarrow{\text{step 1}} \\
 \xrightarrow{\text{step 1}} & |\text{neutron \& meas. dev. \& retina}\rangle = \\
 & = \sum_{j=1}^{87} c_j |x_j\rangle |M_j\rangle |R_j\rangle \quad (9.3c)
 \end{aligned}$$

According to von Neumann, there is also no reason for the retinal state vector to collapse. Rather, the retinal state is transmitted to the brain via nerves. Since the brain consists of atoms, it too is described by quantum theory. We denote the state vectors of the experimenter's consciousness as $|C_j\rangle$. Before the neural signals from his retina reach his consciousness, the consciousness is in the state $|C_0\rangle$. When the consciousness perceives the state of the retina, its state vector becomes entangled with the state vector of the retina:

$$\begin{aligned}
 & \left. \begin{aligned}
 |\text{neutron \& meas. dev. \& retina}\rangle & \stackrel{(9.3c)}{=} \sum_{j=1}^{87} c_j |x_j\rangle |M_j\rangle |R_j\rangle \\
 |\text{consciousness}\rangle & = |C_0\rangle
 \end{aligned} \right\} \\
 \xrightarrow{\text{step 1}} & |\text{neutron \& meas. dev. \& retina \& consciousness}\rangle = \\
 & = \sum_{j=1}^{87} c_j |x_j\rangle |M_j\rangle |R_j\rangle |C_j\rangle \quad (9.3d)
 \end{aligned}$$

At the consciousness, the chain of entanglements ends according to von Neumann's conviction. Now the collapse occurs:

$$\begin{aligned}
& |\text{neutron \& meas. dev. \& retina \& consciousness}\rangle \stackrel{(9.3d)}{=} \\
& = \sum_{j=1}^{87} c_j |x_j\rangle |M_j\rangle |R_j\rangle |C_j\rangle \xrightarrow{\text{step 2}} \\
& \xrightarrow{\text{step 2}} |x_k\rangle |M_k\rangle |R_k\rangle |C_k\rangle \tag{9.3e}
\end{aligned}$$

This can also be written as

$$c_j \xrightarrow{\text{step 2}} \begin{cases} c_j = 1 & \text{if } j = k \\ c_j = 0 & \text{if } j \neq k . \end{cases}$$

The probability that the entangled state vector (9.3d) collapses onto $|x_k\rangle |M_k\rangle |R_k\rangle |C_k\rangle$, but not onto any other vector, is according to Born's rule

$$W(k) \stackrel{(5.18)}{=} |c_k|^2 .$$

The collapse (9.3b) is according to v. Neumann merely an abbreviation for the actual collapse (9.3e).

Thus the collapse happens according to v. Neumann in the consciousness. Why just there? To explain that, v. Neumann states "that measurement or the related process of subjective perception is a new entity relative to the physical environment, and is not reducible to the latter. [...] Nevertheless, it is a fundamental requirement of the scientific viewpoint — the so-called principle of psycho-physical parallelism — that it must be possible so to describe the extra-physical process of subjective perception as if it were in the reality of the physical world; i.e., to assign to its parts equivalent physical processes in the objective environment, in ordinary space." [68, page 272]

Let's try to translate v. Neumann's sentences into plain English: The measurement and its result (in our example: the neutron is detected at position x_{31}) is something that occurs in the "physical

environment”. But the experimenter’s perception of the result — that is, the process in his consciousness — is “a new entity relative to the physical environment, and is not reducible to the latter.” The collapse takes place in consciousness. And according to von Neumann, consciousness exists outside the world comprehensible to natural science; the laws of physics do not apply to consciousness.

Since according to the “principle of psycho-physical parallelism” it must be possible, however, to describe consciousness *as if* it were taking place in the physical world, von Neumann assigns the state vectors $|C_j\rangle$ to it. It is an empirical fact that no experimenter has ever observed a properly functioning measuring device in the “undecided” state $\sum_j c_j |M_j\rangle$, but always in a specific state $|M_k\rangle$. v. Neumann puts it this way: “experience only makes statements of this type: ‘An observer has made a certain (subjective) observation,’ and never any like this: ‘A physical quantity has a certain value.’” [68, page 273] In plain English: One simply has to accept that things are the way they are. Consciousness — however it works — causes entangled state vectors to collapse. This cannot be explained, because consciousness lies beyond the reach of physical laws.

Do the laws of physics really not apply to consciousness?

9.3 Many Worlds

Hugh Everett (1930–1982) and Bryce Seligman DeWitt (1923–2004) developed in the 1950s and 1960s an interpretation of quantum phenomena, in which consciousness is also regarded as part of the physically observable world. Furthermore, they assumed that *all* components of the world — including measuring instruments, humans, and their consciousness — must be described using the methods of quantum theory.

Consequently they concluded, that in the example of measuring the position of a neutron — which was already used in the pre-

vious section — the state vectors of the neutron, the measuring device, and so on, and finally the state vector of the experimenter's consciousness, become entangled with one another:

$$\begin{array}{l}
 |\text{neutron}\rangle = \sum_{j=1}^{87} c_j |x_j\rangle \\
 |\text{meas.dev.}\rangle = |M_0\rangle \\
 |\text{retina}\rangle = |R_0\rangle \\
 |\text{consciousness}\rangle = |C_0\rangle
 \end{array}
 \left. \vphantom{\begin{array}{l} |\text{neutron}\rangle \\ |\text{meas.dev.}\rangle \\ |\text{retina}\rangle \\ |\text{consciousness}\rangle \end{array}} \right\} \xrightarrow{\text{step 1}}$$

$$\xrightarrow{\text{step 1}} |\text{neutron \& meas.dev. \& retina \& consciousness}\rangle =$$

$$= \sum_{j=1}^{87} c_j |x_j\rangle |M_j\rangle |R_j\rangle |C_j\rangle \quad (9.4)$$

v. Neumann had assumed that consciousness is not part of the physical world, but had — in accordance with his “principle of psychophysical parallelism” — attributed state vectors $|C_j\rangle$ to it *as if* consciousness were a component of the physical world. According to Everett and DeWitt, however, consciousness is a perfectly normal component of the physical world, just like neutrons, atoms, and measuring instruments.

This has the advantage that Everett and DeWitt can avoid the obscure “psychophysical parallelism”. But it has the disadvantage that they now needed to explain why the experimenter, when looking at the display of the measuring instrument, perceives a specific measured value M_k , but not a diffuse superposition $\sum_j c_j M_j$ of different measured values. Is there a collapse of the entangled state vector (9.4)? When and why?

Everett and DeWitt presented a surprising answer: they postulated that the state vector of consciousness collapses just as little as the state vectors of neutrons, atoms, and measuring devices. In their interpretation, there is never and nowhere a collapse of state vectors.

Why, then, do we observe a specific measurement result rather than a superposition of different measurement results? According to Everett and DeWitt, we do indeed perceive different measurement results, but we perceive them through different components of our consciousness that cannot communicate with one another. With component C_1 of his consciousness, the experimenter perceives the measurement result M_1 ; with component C_2 of his consciousness, he perceives the result M_2 , and so on; and with component C_{87} of his consciousness, he perceives the result M_{87} .

If I perceive the result M_{12} with the component C_{12} of my consciousness, why don't I notice that the component C_{53} of my consciousness perceives the result M_{53} , which is different from the result M_{12} ? In order for different components of consciousness to learn about each other, they must communicate with one another. In quantum theory, the communication between the components of consciousness C_{12} and C_{53} is described by inserting a specific function C into the projection amplitude of $|C_{12}\rangle$ onto $|C_{53}\rangle$:

$$\text{communication} = \langle C_{53} | C | C_{12} \rangle \stackrel{?}{\neq} 0 \quad (9.5a)$$

Communication can only take place if this projection amplitude is nonzero. If this projection amplitude is zero, then communication between my consciousness components C_{12} and C_{53} is impossible.

Now we must remember that (9.5a) is incomplete. The state vectors C_{12} and C_{53} of my consciousness do not have an independent existence; rather, they exist only as components of the entangled state vector |9.4). Therefore, instead of (9.5a), we must use the complete projection amplitude:

$$\begin{aligned} \text{communication} &= \langle 9.4 | C | 9.4 \rangle = \\ &= \sum_{j=1}^{87} \sum_{n=1}^{87} c_j^* c_n \langle C_j | \langle R_j | \langle M_j | \langle x_j | C | x_n \rangle | M_n \rangle | R_n \rangle | C_n \rangle \stackrel{?}{\neq} 0 \end{aligned} \quad (9.5b)$$

Because the state vector (9.4) appears twice in this projection amplitude, I have renamed the index j to n in one of them, so that the two vectors can be distinguished from one another.

A closer examination of this projection amplitude reveals

$$\langle C_j | \langle R_j | \langle M_j | \langle x_j | C | x_n \rangle | M_n \rangle | R_n \rangle | C_n \rangle \stackrel{\text{FAPP}}{=} 0 \text{ if } j \neq n . \quad (9.6)$$

The acronym FAPP has been introduced by Bell [69] (who also discovered Bell's inequality). It means "for all practical purpose". This means: The mixed terms with $j \neq n$ are not necessarily zero; there is no law of nature that precludes communication between components of consciousness that have perceived different measurement results. But in practice, this communication will never succeed. In practice, only terms with $j = n$ in (9.5b) are nonzero. Each component of consciousness can communicate only with itself, but not with components of consciousness that have perceived different measurement results. In a moment I will explain why this is the case.

(9.6) describes perfect schizophrenia. In the example of determining the position of the neutron, all 87 possible outcomes are actually realized according to Everett and DeWitt: A single neutron is actually detected at all 87 positions x_j , the measuring device actually displays all 87 results M_j simultaneously, all 87 different sensory impressions R_j are actually evoked simultaneously on the experimenter's retina, and trigger all 87 different states of consciousness C_j simultaneously in his brain. But thanks to the perfect schizophrenia (9.6), he notices none of this; instead, each of the 87 components of his consciousness lead him to believe that a single, unambiguous measurement result has occurred.

Every time someone perceives some event, reality splits into all the alternative paths that this event could have taken, and no communication is possible between these alternative realities, which, according to Everett and DeWitt, are all equally real. It has

become customary to refer to the many realities as many worlds, and to Everett and DeWitt's interpretation of quantum phenomena as the Many-Worlds Interpretation.

It is high time to provide the proof for (9.6). First, let us consider the case where the communication function C acts only on some, but not all, components of the entangled system (9.4). Consider e.g. a communication function C that acts only on $|M_n\rangle$, $|R_n\rangle$, and $|C_n\rangle$, but not on the vectors $|x_n\rangle$. Then the contribution of the vectors $|x_j\rangle$ and $|x_n\rangle$ to the projection amplitude can be calculated, even if we do not know the exact form of C :

If C does not act onto the vectors $|x_n\rangle$:

$$(9.6) = \underbrace{\langle x_j | | x_n \rangle}_{(5.19c)} \langle C_j | \langle R_j | \langle M_j | C | M_n \rangle | R_n \rangle | C_n \rangle \quad (9.7)$$

$$= \begin{cases} 1 & \text{if } j = n \\ 0 & \text{if } j \neq n \end{cases}$$

As all eigenvectors of a measuring device are orthogonal to one another, communication is possible in this case only if $j = n$, meaning that each of the 87 components of consciousness can communicate only with itself. Two components of consciousness that have perceived different measurement results can never communicate with one another.

We get the same result if C acts not on the measuring device, or not on the retina, or not on consciousness, but only on the other components of the entangled system (9.4). This is because the various state vectors of these subsystems are also orthogonal to one another.

Terms with $j \neq n$ in the projection amplitude (9.6) can only be nonzero if C acts appropriately on *all* components of the entangled system (9.4) at the same time. It is not enough to simply write down a function C that satisfies this condition. After all, C is merely the mathematical representation of a process that is actually supposed

to take place “out there”, e. g. the transmission and reception of a radio signal, or the exchange of electrical signals between different neurons in the observer’s brain, etc. So, if the communication attempt is to have any chance of success, one must intervene quite aggressively into the observer’s brain. The observer will hardly survive that.

Actually the situation is even *much more* complicated than described so far. After all, the observer is not isolated from the environment, and if he is to survive the measurement process, he cannot be isolated from the environment. He cannot be placed in a vacuum chamber; instead, he must permanently breathe air. Thus, he inevitably becomes entangled with the air molecules in the laboratory, and these, in turn, become entangled with the walls of the laboratory onto which they collide. Furthermore, one cannot cool the observer down to near absolute zero. Thus, he inevitably exchanges uncontrolled infrared thermal radiation with the environment, resulting in further entanglement. Every photon from the laboratory lighting, and every photon of the ubiquitous cosmic background radiation that is scattered by the observer, becomes entangled with him as well. So far, we have neglected to include all these myriads of quantum objects in the state vector (9.4), even though they, too, are components of the entangled quantum system.

If we represent the myriads of molecules and photons in the environment by the state vector $|\text{environment}\rangle = |En\rangle$, then the actual situation is not described by (9.4) and (9.5b), but rather by

$$\left. \begin{aligned} |\text{neutron}\rangle &= \sum_{j=1}^{87} c_j |x_j\rangle \\ |\text{meas.dev.}\rangle &= |M_0\rangle \\ |\text{retina}\rangle &= |R_0\rangle \\ |\text{consciousness}\rangle &= |C_0\rangle \\ |\text{environment}\rangle &= |En_0\rangle \end{aligned} \right\} \xrightarrow{\text{step 1}}$$

step 1 \rightarrow

$$\begin{aligned}
 & |\text{neutron \& meas.dev. \& retina \& consciousness \& environment}\rangle = \\
 & = \sum_{j=1}^{87} c_j |x_j\rangle |M_j\rangle |R_j\rangle |C_j\rangle |En_j\rangle \quad , \quad (9.8)
 \end{aligned}$$

and

$$\begin{aligned}
 \text{communication} &= \langle 9.8 | C | 9.8 \rangle = \sum_{j=1}^{87} \sum_{n=1}^{87} c_j^* c_n \cdot \\
 & \cdot \underbrace{\langle En_j | | En_n \rangle}_{(5.19c)} \langle C_j | \langle R_j | \langle M_j | \langle x_j | C | x_n \rangle | M_n \rangle | R_n \rangle | C_n \rangle \stackrel{?}{\neq} 0 \quad .
 \end{aligned} \quad (9.9)$$

The sheer number of molecules and photons in the environment that are entangled with the observer will make it impossible even for future experimenters, no matter how advanced their technical capabilities may be, to construct a communication mechanism capable of even approximately controlling the entangled quantum system. Therefore, all terms in (9.9) with $j \neq n$ are FAPP zero.

But couldn't the human observer be replaced by a computer built as simple as possible, one that automatically controls the experiment, records the results, and which could — with sufficient technical effort — be isolated from the environment? That could be done, but then nothing would take place that deserves the name “measurement”. For the very essence and purpose of a measurement lies precisely in the fact that a result comes to the attention of human beings, who can evaluate it and discuss it with others.

The inevitable and uncontrollable entanglement of living beings with the environment is important not only for the Many-Worlds Interpretation, but also for the Copenhagen Interpretation. I will return to this consideration in section 9.4.

The Many-Worlds Interpretation is so bizarre that it at first takes one's breath away. Nevertheless, it cannot simply be dismissed as nonsense. On the contrary, this interpretation has several significant advantages. The most important one is certainly that the entire physical reality is described uniformly using a single theory. There is no need to constantly combine classical and quantum theory, as is the case with the Copenhagen Interpretation. Furthermore, as outlandish as it may seem, the Many-Worlds Interpretation is actually very simple. The simplicity of an explanation is consistently recognized by physicists as a strong argument.

Regarding the question of whether humans can choose this or that course of action by virtue of free will, or whether all events in this world — including all supposedly free human decisions — are in fact deterministically predetermined, the Many-Worlds Interpretation offers a peculiar answer: Decisions are neither free nor deterministically predetermined; rather, there are no decisions at all! If I now consider getting up from my desk or staying seated, and decide by virtue of free will to stay seated, then from the perspective of the Many-Worlds Interpretation, this is merely a naive illusion of my consciousness component $|C_{\text{stay-seated}}\rangle$. But in fact, according to Everett and DeWitt, I simultaneously observe with the component $|C_{\text{stand-up}}\rangle$ of my consciousness that I have decided to stand up, and that I have also put this decision into action.

In section 7.5 we had arrived at this conclusion:

- * If the assumptions $A2_{\text{Peres}}$, $A3_{\text{Peres}}$, $A4_{\text{Peres}}$ all three are correct, then the violation of Bell's inequality proves that the result of the measurement is not determined by a yet unknown law of nature; it proves, that the creation of the measurement result in the moment of measurement is really an **irrational** act of Nature which will forever remain beyond the reach of scientific analysis, that genuine chance is at work in the

generation of the measurement result.

Nature's irrational choice of a specific outcome in the moment of measurement is the only(!) instance in physics which is *not* deterministic. This conclusion explicitly presupposes the assumption

A4_{Peres} : Measurements have unambiguous outcomes.

which is no longer true in the Many-Worlds Interpretation. In the Many-Worlds scenario, Nature does not choose a specific outcome, but instead realizes all possible measurement outcomes simultaneously in the various branches of reality. According to the Many-Worlds Interpretation, we live in a world (resp. simultaneously in many worlds) in which no decision is ever made. And if no decision is ever made, then the question of whether decisions are deterministically predetermined or whether humans are free in their decisions is obviously obsolete.

9.4 Decoherence

Figure 9.2 shows four wave trains A1, A2, B1, and B2. If wave trains A1 and A2 are superimposed in an interference experiment, they will nearly cancel out each other, because the waves are out of phase by almost exactly half a wavelength. In contrast, no interference patterns will be observed when the wave trains B1 and B2 are superimposed. This is because there is no clear phase relationship even among the component waves of B1: The phases

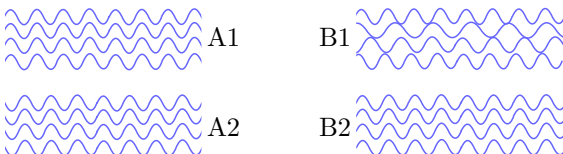


Fig. 9.2: Decoherence

(wave maxima and minima) of the component waves are chaotically shifted relative to one another, and furthermore, the component waves have different wavelengths.

If there is a fixed phase relationship between different wave trains, or between the component waves of a single wave train (e.g., “the maxima of one wave train are shifted by 0.27 wavelengths relative to the maxima of the other wave train”, or similar), then clear interference patterns can be expected when these waves are superimposed. Such waves are referred to as *coherent*.

If there is no fixed phase relationship between two wave trains, or within one or both wave trains, the waves are called *incoherent*. In fig. 9.2, wave train B1 is already incoherent in itself, and therefore also incoherent relative to wave train B2. When originally coherent waves lose their coherence for whatever reason, this process is called *decoherence*.

In section 8.1, we discussed the possibility of narrowing the trajectory of electrons by scattering photons off them. For example, in the double-slit experiment, one could place behind one of the slits a vertically directed beam of light that is so intense that the electron will significantly deflect at least one photon from its path when it passes through that slit. This restricts the electron's trajectory to one of the two slits: If a photon is significantly scattered, then the existence of the scattered photon confirms that the electron passed through that slit. If the electron reaches the detector without a photon being scattered, then it must have passed through the other slit. If the electron scatters a photon, then the electron's momentum changes due to the collision, and consequently its de Broglie wavelength

$$\lambda \stackrel{(4.2b)}{=} \frac{h}{p} = \frac{h}{\text{momentum}}$$

changes by an unpredictable amount. This results in exactly the state illustrated in the right sketch in fig. 9.2: The interaction with

the light barrier behind one of the slits leads to decoherence of the electron wave, causing the interference pattern in the electron detector plane to disappear.

The decoherence of waves causes interference patterns to disappear. In the language of physics, however, the term “decoherence” is defined much more broadly today: *Whenever* interference patterns disappear, this is referred to as decoherence, even if the disappearance of the interference has a completely different cause than a change of wavelength or phase relationship.

For example, in the experiment by Dürr et al. using rubidium atoms, which I have described in section 8.2, the interference disappeared because the atoms were labeled differently as they passed through the interferometer via different paths. I emphasized that this type of labeling had no significant influence on the wavelength and phase of the atom waves. Rather, the interference disappeared, because in (8.4) projection amplitudes of the form

$$\langle 2||3 \rangle = \langle 3||2 \rangle = 0$$

showed up, which have absolutely nothing to do with the wavelength and phase of the de Broglie-wave of the rubidium atoms. Nevertheless, even in this case the disappearance of interference is referred to as “decoherence of the rubidium atoms”.

Exactly the same effect as the deliberate marking of quantum objects has the unintended but, under certain circumstances, unavoidable entanglement of an object with the photons and molecules in its environment. At the end of chapter 4 I wrote: “There is no fundamental reason why interference experiments should not be possible with objects of arbitrarily large mass (for example, with trucks). It is ‘merely’ a matter of the experimenters’ skill in overcoming the technical difficulties.” With today’s state of technology, experimenters cannot prevent an object as large as a truck from being localized by the exchange of thermal radiation with

the environment, nor by air molecules that are **scattered** off these objects. Even in a very good vacuum, far too many molecules of the residual gas are scattered by such a large object for interference experiments to succeed.

If we denote the state vector of the truck as $|T\rangle$, and the state vector of the photons and molecules scattered by the truck as $|En\rangle$, then the state vector of the entire system, when the truck passes through the double slit in the interference experiment, becomes

$$\frac{1}{2} \left(|T\rangle_{\text{left}} |En\rangle_{\text{left}} + |T\rangle_{\text{right}} |En\rangle_{\text{right}} \right). \quad (9.10)$$

The entanglement of the truck with the molecules and photons of the environment has exactly the same effect as the labeling of the rubidium atoms due to the state vectors $|2\rangle$ and $|3\rangle$. Just as the trajectory of the rubidium atoms was no longer delocalized across paths A *and* B of the interferometer, but was restricted to path A *or* B by the labeling with $|2\rangle$ or $|3\rangle$, so the trajectory of the truck is no longer delocalized across both slits, but is restricted to the left or right slit due to the molecules and photons of the environment scattered at the right or left slit. Consequently, projection amplitudes of the form

$$\text{right}\langle En || En \rangle_{\text{left}} = \text{left}\langle En || En \rangle_{\text{right}} = 0$$

prevent the self-interference of the truck, just as in (8.4) projection amplitudes of the form

$$\langle 2 || 3 \rangle = \langle 3 || 2 \rangle = 0$$

prevented the self-interference of the labeled rubidium atom.

This *uncontrolled entanglement* of an object's state vector with the state vectors of molecules and photons of the environment is also referred to as decoherence.⁸⁰ This type of decoherence causes

⁸⁰ Physicists can find a relatively simple introduction to the mathematical formalism of decoherence in [70].

all objects in our everyday environment to appear “classical”. They are incapable of interference, neither with themselves nor with other objects. Their location is precisely defined at any time because it is “measured” — i. e. *created by measurement* — at any time through interaction with countless molecules and photons of the environment.

This brings us to the additional argument of the Copenhagen Interpretation — already announced in section 9.1 — that a classical description of measuring instruments is more adequate to the objective reality than a quantum-theoretical description. Measuring instruments are not only inevitably and uncontrollably entangled with the environment due to their size and complexity; they *must* also be entangled with the environment so that the observer can learn the result of the measurement. A measuring device that were — in order to prevent uncontrolled entanglement with the environment — perfectly isolated from the environment would be completely useless. Measuring devices are therefore never “coherent” quantum objects capable of interference, but are necessarily always precisely localized classical objects. In his 1955 lecture [65] on the Copenhagen Interpretation of Quantum Theory, Heisenberg explains:

“The measuring device deserves this name only if it is in close contact with the rest of the world, if there is an interaction between the device and the observer. [...] If the measuring device would be isolated from the rest of the world, it would be neither a measuring device nor could it be described in the terms of classical physics at all. [...] Therefore, the uncertainty with respect to the microscopic behavior of the entire world will enter into the quantum-theoretical system here.⁸¹ [...] We may say that the transition from

⁸¹ Heisenberg delivered this lecture to a lay audience, whom he did not want to intimidate with mathematical terms. For this reason, he does not explicitly

the ‘possible’ to the ‘actual’ takes place as soon as the interaction of the object with the measuring device, and thereby with the rest of the world, has come into play; it is not connected with the act of registration of the result by the mind of the observer.”

Note that in the last sentence, Heisenberg unequivocally distances himself from v. Neumann’s “collapse” interpretation, which I described in section 9.2. According to the Copenhagen Interpretation (i. e., according to Bohr and Heisenberg), uncontrolled entanglement with the environment — which is practically unavoidable for a large part of the world and even necessary for measuring instruments and for the people who take note of the measurement results — means that this part of the world is correctly described by classical physics. Not only after a “collapse” of the state vector, but from the very beginning!

A description of measuring instruments and human beings as quantum objects, as e. g. v. Neumann attempted to do in (9.3d), is according to this understanding not only clumsy but downright wrong and misleading, because it ignores the uncontrolled entanglement of the devices and people with the environment, even though this entanglement is absolutely inevitable and fundamentally shapes the reality of measuring instruments and human beings.

The classical description of that part of the world which is uncontrollably entangled with its environment is therefore by the Copenhagen Interpretation not regarded as a stopgap measure necessitated by the limited cognitive capabilities of the human brain, but rather as the *only* type of description that is truly appropriate with regard to objective reality.

With electrons, neutrons, atoms, and not too large molecules,

refer to “uncontrolled entanglement with the environment”, but instead vaguely invokes “the uncertainty of the microscopic structure of the entire world.”

interference experiments have been successfully performed. Interference experiments with objects of visible size — or even trucks — have failed so far due to uncontrolled entanglement with the surrounding environment. It would be interesting to investigate precisely, in the intermediate range, how decoherence sets in and how it works. Such an experiment was conducted in 2003 at the University of Vienna by Lucia Hackermüller et al. [71]:

Fig. 9.3 shows the model of a C_{70} -molecule. A beam of C_{70} molecules, whose de Broglie wavelength (4.2b) was approximately 0.0026 nm, was directed onto 3 gold grids as sketched in fig. 9.4 on the facing page. The distance (center to center) between the slits in the grids was 991 nm. The distance from the first to the second and from the second to the third grid was 38 cm. Grid G3 could be moved perpendicular to the direction of motion of the molecules, as indicated by the broad arrows. Behind G3, the C_{70} -molecules were ionized with a strong laser beam, so that the charged parts could be easily detected. The detector counted the molecules that arrived behind grid G3, without distinguishing where the molecules passed through G3.

One does intuitively sense that interference will occur in this experiment. The setup differs significantly, however, from the

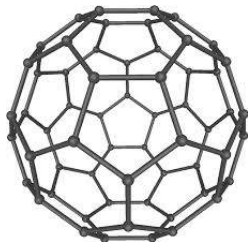


fig. 9.3: A C_{70} molecule.

Each small sphere represents a carbon atom. The atoms are arranged on the surface of an ellipsoid in 25 hexagons and 12 pentagons such that each atom has three nearest neighbors at a distance of about 0.14 nm. The architect Richard Buckminster Fuller (1895–1983) used similar structures in the construction of “geodesic domes”. Therefore spherical carbon molecules of this type are referred to as *fullerenes* or — even more jovial — as *Bucky-Balls*.

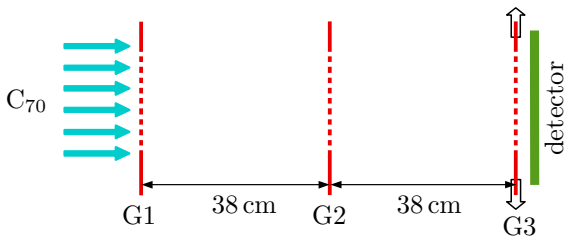


Fig. 9.4: The experiment of Hackermüller et al. [71]

double-slit experiments we examined in sections 4.3 through 4.5. In the experiments described there, so-called far-field interference was investigated, whereas the experiment by Hackermüller et al. deals with near-field interference. The theory of near-field interference, discovered in 1836 by William Henry Fox Talbot (1800–1877), is more complicated and harder to understand than the theory of far-field interference. But for the discussion of decoherence, which is our actual topic at the moment, it suffices to note that interference does indeed occur, as shown in the left diagram of fig. 9.5 on the next page. In this diagram, the red dots indicate the number of C_{70} molecules counted behind G_3 per 2 seconds at different lateral displacements of G_3 .

In Talbot-interference, the distance from one interference maximum to the next is exactly equal to the grid period, i.e. 991 nm. It seems as though the wavelength of the molecules plays no role, different from far-field interference. This is not true, however, because the distance

$$\frac{(\text{grid constant})^2}{\text{wavelength}} = \frac{(991 \text{ nm})^2}{0.0026 \text{ nm}} \approx 38 \text{ cm}$$

from grid to grid was carefully tuned to the de Broglie wavelength of the molecules, as specified in the theory of Talbot-interference.

The yellow line was calculated by the theory of Talbot-interference, and adjusted so that the red points lie on average as closely

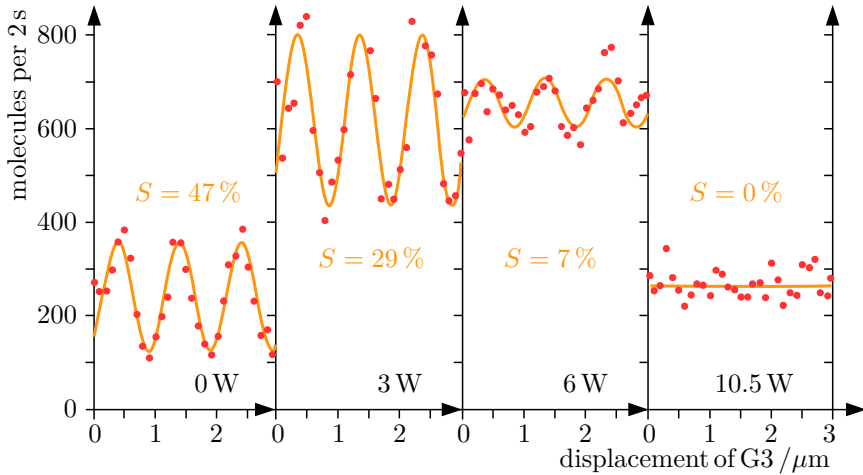


Fig. 9.5: Talbot-Interference of C_{70} -molecules

as possible on it. In addition, the visibility

$$\text{visibility} = S = \frac{\text{count-rate}_{\text{maximum}} - \text{count-rate}_{\text{minimum}}}{\text{count-rate}_{\text{maximum}} + \text{count-rate}_{\text{minimum}}}$$

of the interference pattern as calculated from the yellow curve (not from the red points) is shown in yellow text in the diagram.

The three diagrams on the right in fig. 9.5 show the results from experimental runs in which C_{70} molecules passed immediately before the first grid G1 up to 16 times through a laser beam, which was reflected back and forth across their direction of motion. Each photon absorbed by the molecules heated them by about 150 degrees, i. e. it excited vibrational oscillations of the Bucky-balls (sketched in fig. 9.3). The power of the heating laser is shown in the bottom of the diagrams: 0 Watt, 3 Watt, 6 Watt, 10.5 Watt.

When heated with 3 W and with 6 W, the average count-rate is two to three times higher than at 0 W. The experimenters explain this by noting that the detection of molecules behind G3 is more

efficient because the strongly vibrating molecules can more easily be ionized. Then it is puzzling, however, why the average count-rate is significantly lower again at a heating power of 10.5 W. In any case, the important result of this experiment is the gradual disappearance of the visibility S of the interference as the heating power increases.

The de Broglie wavelength of the C_{70} molecules remains virtually unchanged by heating; this is not the cause of the disappearance of interference. Rather, the disappearance of interference can be explained by the fact that the hot molecules, as they travel between the grids, emit part of their vibrational energy as infrared photons. These infrared photons cause the trajectory of the C_{70} molecules to become more precisely localized, because one could detect these photons and analyze the direction from which they are coming.

Now we can see the similarity to the quantum eraser experiment, which I described in section 8.3. In that experiment, the point of origin of the SPDC-generated photons was localized by the outgoing photon₂ or photon'₂. Now it are the emitted infrared photons that localize the trajectory of the C_{70} molecule.

Different from the case of SPDC-generated photons, however, a single infrared photon is not sufficient to precisely determine the trajectory of a C_{70} molecule. This is because the wavelength of the infrared photons is much longer than the distance (991 nm) between the slits in the grids. But when a large number of infrared photons are emitted, then their combined effect determines the trajectory of the C_{70} molecule so precisely that the interference disappears.

Therefore the disappearance of the interference in this case is a gradual process. The hotter a C_{70} molecule is, the more infrared photons it emits while traveling between the grids, and the more precisely it's trajectory is localized. The localization of the trajectory is an objective fact. It is not necessary that anybody

detects and analyzes the photons. The interference disappears even if the infrared photons escape unrecorded, as in the experiment by Hackermüller et al.

At the end of section 8.3 on the quantum eraser, I emphasized “that the permanent loss of a photon is, in a certain sense, equivalent to its detection by a measuring device. Detection and permanent loss are indeed equivalent in that neither can be reversed.” This is an important point: **It is the defining characteristic of a measurement that it constitutes an *irreversible* process.** No irreversible process, for example, is the passage (transmission or reflection) of a photon through a beam splitter. By simply placing mirrors behind the two outputs of the beam splitter, the process could easily be reversed. In contrast, the detection of a photon using a diode detector is irreversible. The photon triggers an avalanche of electrons in the diode, and no experimenter is capable of reversing the course of this process. In other words: No experimenter is capable of constructing a quantum eraser with which the electrons of the avalanche in the diode could somehow be captured and appropriately redirected for an interference experiment.

It is the decoherence (= uncontrolled entanglement with the environment) of the measuring device that makes a measurement result irreversible. Now, the skills of experimenters have made considerable progress over the past decades and centuries, and we can hope that there will be further significant advances in the coming decades and centuries. Could it be that future experimenters will one day be able to reverse the process of a photon being absorbed by a diode — including the resulting avalanche of electrons — and make it run in reverse? Even if this by far exceeds the capabilities of today’s technology, there is no fundamental reason why it should not be possible in the future.

Then an additional, downstream measuring device would be needed, however, so that the observer can be informed at all about

the process. Because — to paraphrase Heisenberg’s quote printed on page 270 ff — as the diode is now isolated from the rest of the world to prevent decoherence, it cannot be called anymore a “measuring device”. A device is rightly called a measuring device only if it is in close contact with the rest of the world, and can (and must!) therefore be described using the concepts of classical physics.

I have pointed out repeatedly in this book that when measuring a quantum phenomenon, the experimenter does not simply observe an objective reality, but rather creates and shapes this reality — at least in part — through the selection and arrangement of his measuring instruments. Now we see that what constitutes a measurement and what is merely a reversible process is not set in stone but depends on the skills of the experimenters. The ideal of “objectivity”, which was considered sacrosanct in classical physics, thus recedes even further into distance when dealing with quantum phenomena.

9.5 Many Interpretations

Roughly a dozen different interpretations of quantum phenomena have been proposed over the decades. It is impossible to give an exact number, because there is no clear criterion for determining when a modification of a known interpretation should be regarded as a new interpretation, or merely as a variant of the old one. An only halfway complete description and evaluation of the various interpretations would fill a book of its own, one that would be *at least* as thick as this one.⁸²

One cannot say that one interpretation is more correct or incor-

⁸² A brief overview of the most common interpretations can be found on Wikipedia:

https://en.wikipedia.org/wiki/Interpretations_of_quantum_mechanics

rect than another. For in physics, whether something is correct or incorrect is determined by experiments. If a hypothesis has been experimentally disproved, then all physicists agree that the hypothesis was wrong. The interpretations of quantum phenomena, on the other hand, cannot be proved or disproved by experiment. This is because all interpretations are based on the same formal apparatus, on the same mathematical framework used to derive predictions about the statistical frequency of various measurement results according to Born's rule.

What sets these interpretations apart are the vivid images they add to formalism — the background music, so to speak, to the clatter of the mathematical machinery. Is such a thing even necessary? Shouldn't we be satisfied with being able to correctly calculate and predict the results of experiments?

Well, people are different. There are certainly physicists who are completely satisfied with abstract formulas. But others feel that they truly understand a concept only once they have found more or less vivid images that place that concept into a plausible context.

10 Quantum Systems of Many Particles

10.1 Dinge, Bosonen, Fermionen

Wenn man einen starken Lichtstrahl auf einen Strahlteiler richtet, dann wird das Licht zur Hälfte transmittiert, und zur Hälfte reflektiert. Das wissen die Physiker schon seit Jahrhunderten, das wird durch die Klassische Elektrodynamik korrekt beschrieben, das ist leicht zu verstehen. Schwierig und rätselhaft wurde es erst, als die Physiker begannen das Verhalten einzelner Photonen am Strahlteiler zu untersuchen, siehe Abschnitte [3.5](#) und [3.6](#).

Man könnte glauben, dass es sich immer so verhält: Dass die rätselhaften Quantenphänomene auftreten, wenn man sich mit einzelnen Teilchen beschäftigt, aber dass sich die Quantenphänomene irgendwie ausmitteln und zu ganz normalen Klassischen Phänomenen werden, wenn das untersuchte System aus einer sehr großen Anzahl von Teilchen zusammengesetzt ist. Das ist jedoch keineswegs der Fall.

Quantenphänomene mit einzelnen Teilchen gibt es nur in den Laboratorien der Physiker. In technischen Anwendungen hat man es dagegen mit Quantensystemen zu tun, die zum Beispiel aus 10^{20} oder noch mehr Teilchen bestehen. Es ist klar dass man den Zustandsvektor eines so großen Quantensystems unmöglich explizit berechnen kann. Quantensysteme, die aus sehr vielen Teilchen bestehen, können nur mit geeigneten statistischen Methoden beschrieben werden.

Quantenobjekte sind keine Dinge. Quantenobjekte sind etwas

grundlegend anderes, für das es in der menschlichen Sprache keine passenden Begriffe gibt. Der Unterschied tritt auch in der Statistik von Dingen und Quantenobjekten deutlich zutage.

Die Statistik der *Dinge* wird in der Physik als Maxwell-Boltzmann-Statistik bezeichnet, benannt nach James Clerk Maxwell (1831–1879) und Ludwig Boltzmann (1844–1906), die besonders wichtige Beiträge zur Entwicklung dieser Art der Statistik leisteten.

Im Fall von Quantenobjekten hat sich herausgestellt, dass es zwei grundlegend unterschiedliche Arten von Quantenobjekten gibt, die als *Bosonen* und *Fermionen* bezeichnet werden.

Für Bosonen gilt die Bose-Einstein-Statistik⁸³, benannt nach Satyendranath Bose (1894–1974) und Albert Einstein. Für Fermionen gilt die Fermi-Dirac-Statistik, benannt nach Enrico Fermi (1901–1954) und Paul Dirac.

Jedes Quantenobjekt gehört eindeutig entweder zur Gruppe der Bosonen oder zur Gruppe der Fermionen. Beispielsweise sind Elektronen, Protonen, Neutronen, Quarks, und viele Atome und viele Moleküle Fermionen. Beispiele für Bosonen sind Photonen, Gluonen, und viele Atome und viele Moleküle. Auch Schallwellen in Festkörpern, und die Vibrationsschwingungen von Molekülen, sind Bosonen.⁸⁴

Der Unterschied zwischen den drei verschiedenen Statistiken für Dinge, für Bosonen, und für Fermionen lässt sich am einfachsten am Beispiel von zwei Würfeln erklären. Würfel sind Dinge, für sie

⁸³ Kurioserweise wurde die Bose-Einstein-Statistik erstmals bereits im Jahr 1877 in einer Veröffentlichung [72] von Boltzmann angewendet. Boltzmann bezeichnete das aber ausdrücklich als eine lediglich „mathematische Fiktion“. Die wahre Bedeutung dieser Entdeckung erkannten damals weder Boltzmann selbst noch seine Leser. Eine (nur für Physiker geeignete) Darstellung der Entdeckung der Bose-Einstein-Statistik findet man in [73].

⁸⁴ Genau wie Lichtwellen in bestimmten Experimenten als Teilchen (nämlich Photonen) in Erscheinung treten, treten Schallwellen in Festkörpern in bestimmten Experimenten als Teilchen, die Phononen genannt werden, in Erscheinung.

gilt also die Maxwell-Boltzmann-Statistik. Wir nehmen an, dass die beiden Würfel wie üblich auf jeder ihrer sechs Seiten mit ein bis sechs Augen markiert sind. Wenn man beide Würfel gleichzeitig wirft, erhält man Ergebnisse zwischen 2 Augen und 12 Augen, siehe Tabelle 10.1.

Die 36 möglichen Ergebnisse	Augen	Wahrscheinlichkeit
	2	1/36
	3	2/36
	4	3/36
	5	4/36
	6	5/36
	7	6/36
	8	5/36
	9	4/36
	10	3/36
	11	2/36
	12	1/36

Tab. 10.1 : Mögliche Ergebnisse mit zwei Würfeln
(Maxwell-Boltzmann-Statistik)

Jedes einzelne der 36 möglichen Ergebnisse kommt mit gleicher Wahrscheinlichkeit, nämlich $1/36$, vor. Das Ergebnis 7 Augen kann durch sechs verschiedene Kombinationen zustande kommen, das Ergebnis 2 Augen aber nur durch eine einzige Kombination. Also erwartet man, und findet es bei Nachprüfung⁸⁵ auch bestätigt, dass das Ergebnis 7 Augen beim Spiel mit zwei Würfeln sechs mal so häufig vorkommt wie das Ergebnis 2 Augen.

Das Beispielsystem „zwei Würfel“ besteht aus zwei Teilchen (nämlich den beiden Würfeln), die je sechs verschiedene Eigenschaften haben können (nämlich je sechs verschiedene Augenzahlen).

⁸⁵ Wer auch nur den geringsten Zweifel an der Richtigkeit von Tabelle 10.1 hat, sollte das durch ein paar hundert Würfe selbst nachprüfen!

Weil die Würfel Dinge sind, ist $\begin{smallmatrix} \blacksquare & \blacksquare \\ \square & \square \end{smallmatrix}$ ein anderer Zustand als $\begin{smallmatrix} \square & \blacksquare \\ \square & \square \end{smallmatrix}$. Es macht einen Unterschied, ob der weiße Würfel 5 Augen zeigt und der schwarze Würfel 6, oder ob der weiße Würfel 6 Augen zeigt und der schwarze Würfel 5. Deshalb werden in Tabelle 10.1 bei der Berechnung der Wahrscheinlichkeit des Gesamtergebnisses 11 Augen $\begin{smallmatrix} \blacksquare & \blacksquare \\ \square & \square \end{smallmatrix}$ und $\begin{smallmatrix} \square & \blacksquare \\ \square & \square \end{smallmatrix}$ als unterschiedliche Zustände des Systems gezählt, so dass die Wahrscheinlichkeit für 11 Augen gleich $2/36$ ist.

Die Erfahrung zeigt, dass für Quantenobjekte andere statistische Regeln gelten als für Dinge:

- * Wenn in einem System gleichartiger Bosonen die Eigenschaften von zwei Teilchen miteinander vertauscht werden, dann ändert sich der Zustandsvektor des Gesamtsystems um den Faktor $(+1)$, d. h. er bleibt unverändert. Man sagt, dass der Zustandsvektor eines Systems gleichartiger Bosonen unter Vertauschungen *symmetrisch* ist. (10.1a)
- * Wenn in einem System gleichartiger Fermionen die Eigenschaften von zwei Teilchen miteinander vertauscht werden, dann ändert sich der Zustandsvektor des Gesamtsystems um den Faktor (-1) . Man sagt, dass der Zustandsvektor eines Systems gleichartiger Fermionen unter Vertauschungen *antisymmetrisch* ist. (10.1b)

Warum ist das so? Die Physik kennt für (10.1) keine tiefer liegende Begründung. Es handelt sich um ein Naturgesetz, das aus der Analyse von Experimenten erraten wurde, und durch alle experimentellen Beobachtungen bestätigt wird.

Was (10.1) konkret bedeutet versteht man am leichtesten, wenn wir einmal annehmen wir könnten mit zwei Quantenwürfel würfeln, durch zahlreiche Versuche die Wahrscheinlichkeiten der verschiedenen Ergebnisse ermitteln, und in einer Tabelle zusammenstellen.

Tatsächlich gibt es in der Natur keine Quantenwürfel, und bis heute ist kein Experimentator in der Lage, einen Quantenwürfel herzustellen. Trotzdem ist es nützlich einmal zu überlegen, welche Auswirkungen das Naturgesetz (10.1) auf die Ergebnisse beim Würfeln mit Quantenwürfeln hätte. In den Abschnitten 10.2 und 10.3 werden wir dann zwei realistischere Beispiele betrachten.

Wenn ein Würfel fünf Augen zeigt und der andere drei, dann kann der Zustandsvektor eines Quantensystems, das aus zwei Quantenwürfeln besteht, weder $|\otimes\rangle|\otimes\rangle$ noch $|\otimes\rangle|\otimes\rangle$ sein. Denn wenn die Augenzahlen der beiden Würfel miteinander vertauscht werden, dann ändern sich diese Zustandsvektoren weder um den Faktor (+1) noch um den Faktor (-1), sondern dann ändert sich der eine dieser Zustandsvektoren in den anderen, und der andere in den einen.

Der richtige Zustandsvektor ist vielmehr im Fall von

$$\text{Bosonen : } \sqrt{\frac{1}{2}} \left(|\otimes\rangle|\otimes\rangle + |\otimes\rangle|\otimes\rangle \right) \quad (10.2a)$$

$$\text{Fermionen : } \sqrt{\frac{1}{2}} \left(|\otimes\rangle|\otimes\rangle - |\otimes\rangle|\otimes\rangle \right) . \quad (10.2b)$$

Diese Zustandsvektoren haben bei Vertauschung der Augenzahlen der beiden Würfel das richtige Verhalten entsprechend (10.1):

$$\begin{aligned} \left(|\otimes\rangle|\otimes\rangle + |\otimes\rangle|\otimes\rangle \right) & \xrightarrow{\text{Vertauschung}} \left(|\otimes\rangle|\otimes\rangle + |\otimes\rangle|\otimes\rangle \right) = \\ & = + \left(|\otimes\rangle|\otimes\rangle + |\otimes\rangle|\otimes\rangle \right) \end{aligned}$$

$$\begin{aligned} \left(|\otimes\rangle|\otimes\rangle - |\otimes\rangle|\otimes\rangle \right) & \xrightarrow{\text{Vertauschung}} \left(|\otimes\rangle|\otimes\rangle - |\otimes\rangle|\otimes\rangle \right) = \\ & = - \left(|\otimes\rangle|\otimes\rangle - |\otimes\rangle|\otimes\rangle \right) \end{aligned}$$

Der Faktor $\sqrt{1/2}$ in (10.2) ist erforderlich, damit die Projektionsamplitude der Zustandsvektoren auf sich selbst 1 ergibt, wie es sein muss:

$$\begin{aligned}
& \sqrt{\frac{1}{2}} \left(\langle \text{⊠} | \langle \text{⊠} | + \langle \text{⊠} | \langle \text{⊠} | \right) \sqrt{\frac{1}{2}} \left(| \text{⊠} \rangle | \text{⊠} \rangle + | \text{⊠} \rangle | \text{⊠} \rangle \right) = \\
& = \frac{1}{2} \left(\underbrace{\langle \text{⊠} | \langle \text{⊠} |}_{1} \underbrace{| \text{⊠} \rangle | \text{⊠} \rangle}_{1} + \underbrace{\langle \text{⊠} | \langle \text{⊠} |}_{0} \underbrace{| \text{⊠} \rangle | \text{⊠} \rangle}_{0} + \right. \\
& \quad \left. + \underbrace{\langle \text{⊠} | \langle \text{⊠} |}_{0} \underbrace{| \text{⊠} \rangle | \text{⊠} \rangle}_{0} + \underbrace{\langle \text{⊠} | \langle \text{⊠} |}_{1} \underbrace{| \text{⊠} \rangle | \text{⊠} \rangle}_{1} \right) = 1 \quad (10.3a)
\end{aligned}$$

$$\begin{aligned}
& \sqrt{\frac{1}{2}} \left(\langle \text{⊠} | \langle \text{⊠} | - \langle \text{⊠} | \langle \text{⊠} | \right) \sqrt{\frac{1}{2}} \left(| \text{⊠} \rangle | \text{⊠} \rangle - | \text{⊠} \rangle | \text{⊠} \rangle \right) = \\
& = \frac{1}{2} \left(\underbrace{\langle \text{⊠} | \langle \text{⊠} |}_{1} \underbrace{| \text{⊠} \rangle | \text{⊠} \rangle}_{1} - \underbrace{\langle \text{⊠} | \langle \text{⊠} |}_{0} \underbrace{| \text{⊠} \rangle | \text{⊠} \rangle}_{0} - \right. \\
& \quad \left. - \underbrace{\langle \text{⊠} | \langle \text{⊠} |}_{0} \underbrace{| \text{⊠} \rangle | \text{⊠} \rangle}_{0} + \underbrace{\langle \text{⊠} | \langle \text{⊠} |}_{1} \underbrace{| \text{⊠} \rangle | \text{⊠} \rangle}_{1} \right) = 1 \quad (10.3b)
\end{aligned}$$

Falls die beiden Würfel die gleiche Augenzahl („Pasch“) zeigen, beispielsweise zwei Einsen, ergeben die Zustandsvektoren (10.2) im Fall von

$$\text{Bosonen : } \sqrt{\frac{1}{2}} \left(| \text{⊠} \rangle | \text{⊠} \rangle + | \text{⊠} \rangle | \text{⊠} \rangle \right) = | \text{⊠} \rangle | \text{⊠} \rangle \quad (10.4a)$$

$$\text{Fermionen : } \sqrt{\frac{1}{2}} \left(| \text{⊠} \rangle | \text{⊠} \rangle - | \text{⊠} \rangle | \text{⊠} \rangle \right) = 0 . \quad (10.4b)$$

Dass die Faktoren $\sqrt{1/2}$ und 1 im Fall des Bosonen-Paschs richtig sind, sieht man wenn man jeweils die Projektionsamplitude der rechten und der linken Seite von (10.4a) auf sich selbst berechnet. Bemerkenswert ist (10.4b): Weil der Zustandsvektor Null ist, ist seine Projektionsamplitude auf sich selbst erst recht Null, d. h. die Wahrscheinlichkeit eines Paschs ist bei fermionischen Würfeln Null. Fermionische Würfel haben aufgrund des Naturgesetzes (10.1b) *immer* unterschiedliche Augenzahlen.

In Tabelle 10.2 auf der nächsten Seite sind die möglichen Ergebnisse beim Würfeln mit zwei Bosonenwürfeln eingetragen. Es

Die 21 möglichen Ergebnisse	Augen	Wahrscheinlichkeit
	2	1/21
$(\text{white} + \text{black})$	3	1/21
$(\text{white} + \text{black}), (\text{black} + \text{white})$	4	2/21
$(\text{white} + \text{black}), (\text{black} + \text{white})$	5	2/21
$(\text{white} + \text{black}), (\text{black} + \text{white}), (\text{white} + \text{black})$	6	3/21
$(\text{white} + \text{black}), (\text{black} + \text{white}), (\text{black} + \text{white}), (\text{white} + \text{black})$	7	3/21
$(\text{white} + \text{black}), (\text{black} + \text{white}), (\text{black} + \text{white}), (\text{white} + \text{black})$	8	3/21
$(\text{white} + \text{black}), (\text{black} + \text{white}), (\text{black} + \text{white}), (\text{white} + \text{black})$	9	2/21
$(\text{white} + \text{black}), (\text{black} + \text{white}), (\text{black} + \text{white}), (\text{white} + \text{black})$	10	2/21
$(\text{white} + \text{black}), (\text{black} + \text{white}), (\text{black} + \text{white}), (\text{white} + \text{black})$	11	1/21
$(\text{white} + \text{black}), (\text{black} + \text{white}), (\text{black} + \text{white}), (\text{white} + \text{black})$	12	1/21

Tab. 10.2: Mögliche Ergebnisse mit zwei bosonischen Würfeln (Bose-Einstein-Statistik)

gibt nur noch 21 verschiedene Ergebnisse, im Gegensatz zu den 36 möglichen Ergebnissen, wenn es sich bei den Würfeln um Dinge handelt.

In den Zustandsvektoren

$$\underbrace{\sqrt{\frac{1}{2}} \left(|\text{white}| |\text{black}| + |\text{black}| |\text{white}| \right)}_{\text{Bosonen}} \quad \text{bzw.} \quad \underbrace{\sqrt{\frac{1}{2}} \left(|\text{white}| |\text{black}| - |\text{black}| |\text{white}| \right)}_{\text{Fermionen}} \quad (10.5)$$

haben die einzelnen Würfel ihre Identität verloren: Das System besteht eindeutig aus einem weißen und einem schwarzen Würfel. Und es ist eindeutig, dass die Gesamtzahl der Augen 8 ist, zusammengesetzt aus 5 Augen und 3 Augen der einzelnen Würfel. Aber die Eigenschaft 3 Augen bzw. 5 Augen kann nicht eindeutig dem weißen bzw. dem schwarzen Würfel zugeordnet werden.

Diese Eigenart von Quantensystemen in verschränkten Zuständen wurden in Kapitel 6 ausführlich untersucht. Auch (10.5) sind verschränkte Zustandsvektoren, die das Gesamtsystem in ganzheitlicher Weise beschreiben. Das Gesamtsystem hat genau definierte

Eigenschaften (hier: Augenzahlen), aber seine Bestandteile (hier: der weiße und der schwarze Würfel) existieren nur als Bestandteile des Gesamtsystems, haben keine eigenständige Existenz, und deshalb auch keine genau definierten Eigenschaften (hier: Augenzahlen).

Wenn man mit drei fermionischen Würfeln würfelt, dann ist ein mögliches Ergebnis 2 Augen, 3 Augen, 5 Augen. Der Zustandsvektor

$$\sqrt{\frac{1}{6}} \left(|\text{⊠}\rangle|\text{⊞}\rangle|\text{⊚}\rangle - |\text{⊠}\rangle|\text{⊞}\rangle|\text{⊚}\rangle - |\text{⊠}\rangle|\text{⊞}\rangle|\text{⊚}\rangle + |\text{⊞}\rangle|\text{⊞}\rangle|\text{⊚}\rangle + |\text{⊠}\rangle|\text{⊞}\rangle|\text{⊚}\rangle - |\text{⊞}\rangle|\text{⊞}\rangle|\text{⊚}\rangle \right) \quad (10.6)$$

erfüllt die Bedingung 10.1b: Der Zustandsvektor wechselt das Vorzeichen, egal ob man in jedem der sechs Summanden die Augenzahl des weißen und des schwarzen Würfels vertauscht, oder ob man die Augenzahl des weißen und des roten Würfels vertauscht, oder ob man die Augenzahl des schwarzen und des roten Würfels vertauscht. Wer mag kann als Übungsaufgabe nachrechnen, dass die Projektionsamplitude des Zustands (10.6) auf sich selbst 1 ergibt; der Faktor $\sqrt{1/6}$ ist also richtig.

Niemals kann in einem System von fermionischen Würfeln die gleiche Augenzahl zweimal (oder gar noch öfter) vorkommen, denn dann ist der Zustandsvektor Null. Ein Beispiel ist das (unmögliche) Ergebnis 2 Augen, 5 Augen, 5 Augen. Der Zustandsvektor dieses Ergebnisses ist

$$\sqrt{\frac{1}{6}} \left(\underbrace{|\text{⊠}\rangle|\text{⊞}\rangle|\text{⊚}\rangle - |\text{⊠}\rangle|\text{⊞}\rangle|\text{⊚}\rangle}_{0} - \underbrace{|\text{⊞}\rangle|\text{⊞}\rangle|\text{⊚}\rangle + |\text{⊞}\rangle|\text{⊞}\rangle|\text{⊚}\rangle}_{0} + \underbrace{|\text{⊞}\rangle|\text{⊞}\rangle|\text{⊚}\rangle - |\text{⊞}\rangle|\text{⊞}\rangle|\text{⊚}\rangle}_{0} \right) = 0. \quad (10.7)$$

Weil der Zustandsvektor Null ist, ist die Projektionsamplitude dieses Zustandsvektors auf sich selbst erst recht Null. Das bedeu-

Die 15 möglichen Ergebnisse	Augen	Wahrscheinlichkeit
$(\square \blacksquare - \square \blacksquare)$	3	1/15
$(\square \blacksquare - \square \blacksquare)$	4	1/15
$(\square \blacksquare - \square \blacksquare), (\square \blacksquare - \square \blacksquare)$	5	2/15
$(\square \blacksquare - \square \blacksquare), (\square \blacksquare - \square \blacksquare)$	6	2/15
$(\square \blacksquare - \square \blacksquare), (\square \blacksquare - \square \blacksquare), (\square \blacksquare - \square \blacksquare)$	7	3/15
$(\square \blacksquare - \square \blacksquare), (\square \blacksquare - \square \blacksquare)$	8	2/15
$(\square \blacksquare - \square \blacksquare), (\square \blacksquare - \square \blacksquare)$	9	2/15
$(\square \blacksquare - \square \blacksquare)$	10	1/15
$(\square \blacksquare - \square \blacksquare)$	11	1/15

Tab. 10.3: Mögliche Ergebnisse mit zwei fermionischen Würfeln (Fermi-Dirac-Statistik)

tet, dass die Wahrscheinlichkeit des Ergebnisses 2 Augen, 5 Augen, 5 Augen beim Würfeln mit fermionischen Würfeln Null ist, d. h. dass dies Ergebnis unmöglich ist.

Weil beim Würfeln mit fermionischen Würfeln niemals zwei Würfel die gleiche Augenzahl haben können (es gibt kein „Pasch“), bleiben im Fall von zwei Würfeln nur die 15 möglichen Ergebnisse, die in Tabelle 10.3 aufgelistet sind.

Was passiert, wenn man mit sieben fermionischen Würfeln würfelt? Müssen dann nicht mindestens zwei Würfel die gleiche Augenzahl zeigen, weil es ja insgesamt nur die Augenzahlen 1 bis 6 gibt? Nun, ich weiß es auch nicht. Vermutlich ist das Beispiel von fermionischen Quantenwürfeln allzu unrealistisch. Beschäftigen wir uns besser mit realistischen Fermionen-Systemen, zum Beispiel den Elektronen in einem Festkörper.

10.2 Halbleiter-Elektronik

Ein Kristall aus Silizium mit dem Volumen 1 mm^3 enthält etwa $2 \cdot 10^{22}$ Elektronen. Weil Elektronen Fermionen sind, können nicht zwei von ihnen identische Eigenschaften haben! Umgekehrt gesagt:

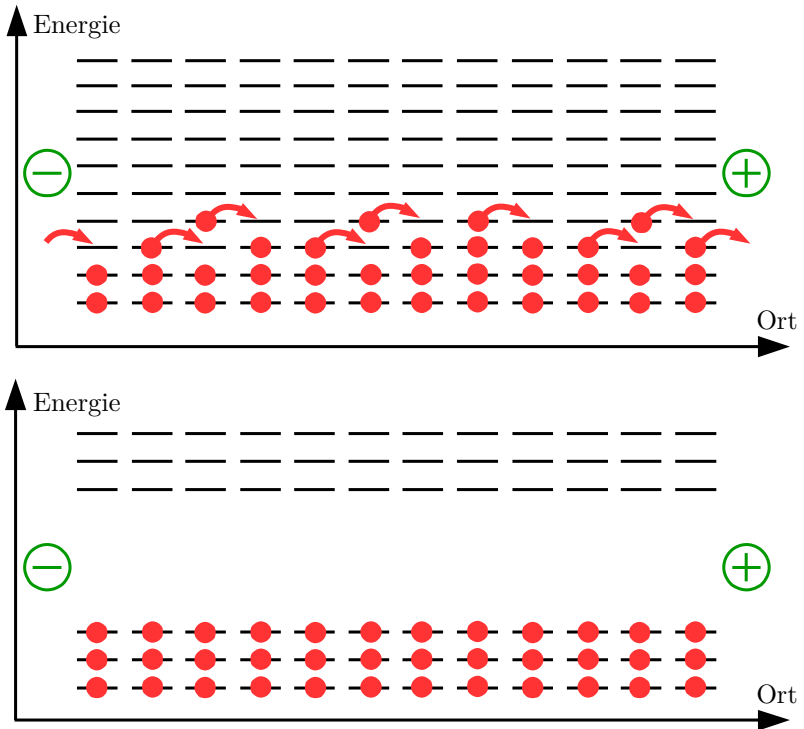


Abb. 10.1: Energieniveaus in einem guten Leiter (oben) und in einem Isolator (unten)

Jedes der $2 \cdot 10^{22}$ Elektronen in diesem Kristall muss sich durch mindestens eine Eigenschaft von jedem anderen der Elektronen unterscheiden.

Unterschiedliche Eigenschaften, das bedeutet konkret z. B. unterschiedlicher Impuls, oder unterschiedlicher Ort, oder unterschiedliche Energie der einzelnen Elektronen. In Abb. 10.1 ist ein extrem vereinfachtes Modell skizziert: Die Elektronen, symbolisiert durch die roten Punkte, haben in diesem Modell nur zwei Eigenschaften, nämlich einen Ort längs einer bestimmten Raumachse und eine

bestimmte Energie. Jedes Elektron in diesem Modellkristall muss sich entweder durch seinen Ort oder durch seine Energie von jedem anderen Elektron unterscheiden, d. h. jeder der kleinen waagerechten Striche, die die möglichen Orte und Energien von Elektronen symbolisieren, darf nur von maximal einem Elektron belegt sein.

Die grün eingekreisten Plus- und Minuszeichen sollen eine elektrische Spannung symbolisieren, die rechts und links an die Oberflächen des Kristalls angelegt wird. Gleichnamige Ladungen stoßen sich ab, ungleichnamige Ladungen ziehen sich an. Elektronen sind negativ geladen, also bewegen sie sich, wie durch die roten Pfeile angedeutet, nach rechts in Richtung zum Pluspol – wenn sie können.

Bewegen können die Elektronen sich nur dann, wenn der Platz zu dem sie sich hinbewegen wollen, nicht bereits belegt ist. In der oberen Skizze gibt es dicht oberhalb der belegten Energieniveaus zahlreiche freie Energieniveaus. Der energetische Abstand zwischen den belegten und den freien Niveaus ist so klein, dass die Elektronen durch thermische Anregung (Energietransfer von den Gitterschwingungen des Kristalls zu den Elektronen) die freien Niveaus leicht erreichen können.⁸⁶ Dann steht ihrer Diffusion durch den Kristall Richtung Pluspol nichts mehr im Weg.

In der unteren Skizze von Abb. 10.1 gibt es zwischen den belegten und den freien Energieniveaus eine große Lücke. Die Energielücke ist viel größer als die thermische Energie. Deshalb wird kein Elektron in die freien Niveaus angeregt. Weil alle erreichbaren Nachbar-Niveaus belegt sind, kann sich keines der Elektronen von der Stelle rühren, dieser Kristall ist ein elektrischer Isolator.

⁸⁶ Tatsächlich ist – anders als in der Skizze – der energetische Abstand zwischen den belegten und den freien Energieniveaus in Metallen so winzig klein, dass er auch mit moderner Messtechnik nicht nachweisbar ist. Das hat zur Folge, dass in Metallen auch bei extremer Kühlung (beliebig nah am absoluten Nullpunkt der Temperatur) Elektronen thermisch in freie Niveaus angeregt werden können.

Aber könnten nicht die Elektronen, die in der unteren Skizze von Abb. 10.1 ganz rechts sitzen, in die (in der Skizze nicht sichtbare) metallische **Elektrode** diffundieren? Dann würden ihre Plätze frei, die nächsten Nachbarlektronen könnten nachrücken, und dann wiederum deren nächste Nachbarn, so dass schließlich doch ein Elektronenstrom durch den Isolator entstehen könnte. Ich werde gleich auf diese Frage zurückkommen; zunächst behaupte ich einfach, dass diese Art von Leitung nicht zustande kommt; der Isolator ist und bleibt ein Isolator.

Die Energielücke zwischen den von Elektronen belegten Energieniveaus und den freien Energieniveaus ist das Markenzeichen von Isolatoren. Das Markenzeichen von Metallen ist, dass bei ihnen zwischen belegten und freien Energieniveaus *keine* Energielücke klappt. Die in der Elektronik-Industrie verwendeten Halbleiter (am häufigsten wird Silizium eingesetzt) tragen ihren Namen zu Unrecht. Sie sind nicht schlecht leitende Metalle, sondern sie sind schlecht isolierende Isolatoren. Halbleiter sind Isolatoren, weil bei ihnen zwischen den belegten und den freien Energieniveaus eine Energielücke von erheblicher Größe besteht. Die Energielücke ist jedoch nicht groß genug, um thermische Anregung von Elektronen vollständig zu unterbinden; dadurch bleibt eine geringe Leitfähigkeit, die aber winzig klein ist im Vergleich zur Leitfähigkeit von Metallen.⁸⁷

Man kann die Leitfähigkeit von Halbleitern deutlich erhöhen, indem man sie mit geeigneten Fremdatomen dotiert. Betrachten wir am Beispiel von Silizium, was das bedeutet. Silizium-Atome haben 14 Elektronen, Aluminium-Atome haben 13 Elektronen, und Phosphor-Atome haben 15 Elektronen. Wenn man einen Silizium-Kristall bei hoher Temperatur mit Strahlen von Al- oder P-Ionen beschießt, dann kann man erreichen dass beispielsweise

⁸⁷ Das gilt bei Raumtemperatur. Wenn man Silizium auf tiefe Temperatur abkühlt, wird es zu einem ausgezeichneten Isolator.

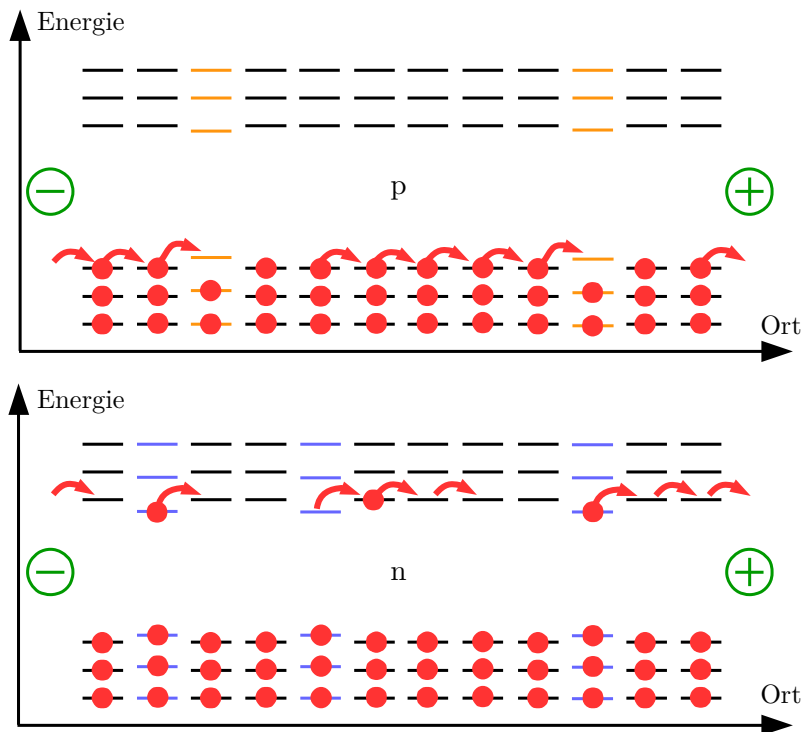


Fig. 10.2: Energieniveaus in p-dotiertem Silizium (oben) und in n-dotiertem Silizium (unten)

durchschnittlich jedes 10 000 ste Si-Atom im Kristall durch ein Al- oder P-Atom ersetzt wird. Die Folgen dieser Dotierung werden in Abb. 10.2 dargestellt. Die obere Skizze zeigt die Energieniveaus eines Aluminium-dotierten Silizium-Kristalls, die untere zeigt die Energieniveaus eines Phosphor-dotierten Silizium-Kristalls. Die Energieniveaus an den Orten der Al-Atome sind in gelber Farbe eingetragen, die Energieniveaus an den Orten der P-Atome in blauer Farbe.

Die Fremdatome versuchen, sich möglichst homogen in den Silizium-Kristall einzufügen. Das gelingt ihnen aber nur unvollkommen, weil sie ja ein Elektron zu viel oder zu wenig haben. Al-Atome haben ein Elektron weniger als Si-Atome. Wenn es einem Al-Atom gelingt, ein zufällig in der Nähe vorbeidiffundierendes Elektron einzufangen, dann kann es sich besser in den Si-Kristall einfügen. Aus diesem Grund ist nur sehr wenig Energie erforderlich, damit ein Elektron aus der Nachbarschaft zum Al-Atom wechselt, d. h. das unterste freie Energieniveau der Al-Atome liegt nur geringfügig höher als die höchsten besetzten Energieniveaus der benachbarten Si-Atome.

Deshalb kann ein Elektron eines Si-Atoms leicht thermisch dazu angeregt werden, zum Al-Atom zu wechseln und dadurch die Kristallstruktur um das Al-Atom herum zu perfektionieren. Wenn das geschieht, dann zieht es eine Kettenreaktion nach sich. Denn dann wird ja der ursprüngliche Platz dieses Elektrons frei, ein Nachbar-elektron kann auf diesen Platz wechseln und macht dabei seinen eigenen Platz frei, und so weiter. Jedes mal, wenn ein Elektron – wie durch die roten Pfeile angedeutet – um einen Platz nach rechts rückt, dann rückt der freie Platz nach links weiter. Die einzelnen Elektronen bewegen sich jeweils nur um einen einzigen Platz nach rechts. Aber der freie Platz, von den Halbleiter-Elektronikern prosaisch als *Loch* bezeichnet, bewegt sich nach und nach von rechts nach links durch den gesamten Kristall.

Im Effekt ist es so, als würde das *Loch* als positiv geladenes Teilchen von rechts nach links durch den Kristall diffundieren. Deshalb wird die Leitung mithilfe von Löchern als p-Leitung bezeichnet, und mit Al dotiertes Silizium als p-dotierter Halbleiter bzw. einfach p-Halbleiter. Das p steht für „positive Ladung“. Tatsächlich existiert positive Ladung nur in den Atomkernen, und die sind fest an ihrem jeweiligen Platz im Kristall eingebaut. Nur die negativ geladenen Elektronen sind beweglich. Das p in p-Halbleiter weist

lediglich darauf hin, dass es sich so verhält *als ob* sich positive Ladungen durch den Kristall bewegen.

P-Atome (nicht das P von Phosphor mit dem p von p-Halbleiter verwechseln!) können sich besser in den Silizium-Kristall einfügen, wenn sie ihr 15. Elektron abgeben (Si-Atome haben nur 14 Elektronen). Deshalb liegt das höchste belegte Energieniveau der P-Atome im Silizium-Kristall dicht unterhalb der ersten freien Energieniveaus der benachbarten Si-Atome, siehe die untere Skizze von Abb. 10.2. Die Energieniveaus der P-Atome sind dort blau dargestellt. Das 15. Elektron eines Phosphor-Atoms kann thermisch leicht in die freien Energieniveaus des Siliziums angeregt werden, und sich dann ungehindert bis zur + Elektrode bewegen. Weil die beweglichen Ladungen (sprich: die Elektronen) in diesem Fall negativ geladen sind, wird diese Art der Leitung als n-Leitung bezeichnet, und mit P dotiertes Silizium als n-dotierter Halbleiter bzw. einfach n-Halbleiter.

Jetzt ist es an der Zeit, auf die Frage mit den Elektroden zurück zu kommen, die ich oben einfach beiseite geschoben hatte: In der unteren Skizze von Abb. 10.2 habe ich durch einen Pfeil ganz links angedeutet, dass Elektronen von der Minus-Elektrode in die freien Energieniveaus des Si-Kristalls eingespeist werden. Das muss so sein, denn sonst wären bald alle 15. Elektronen der P-Atome zur Plus-Elektrode abgeflossen, und dann käme die n-Leitung zum Stillstand. Und in der oberen Skizze habe ich durch einen Pfeil ganz rechts angedeutet, dass Elektronen aus den obersten belegten Energieniveaus des Si-Kristalls in die Plus-Elektrode abfließen können, sprich dass Löcher von der Plus-Elektrode in die obersten belegten Energieniveaus des Si-Kristalls eingespeist werden. Das muss so sein, denn sonst wären bald alle Löcher zur Minus-Elektrode abgeflossen, und dann käme die p-Leitung zum Stillstand.

Wenn die Elektroden Löcher in den p-Halbleiter bzw. Elektronen

in den n-Halbleiter einspeisen, warum speisen sie dann weder Elektronen noch Löcher in den undotierten Si-Kristall von Abb. 10.1 auf Seite 288 ein? Das liegt daran, dass das in den Abbildungen 10.1 und 10.2 skizzierte Modell einen grundlegenden Fehler enthält – wie aufmerksame Leser längst bemerkt haben dürften. Elektronen „haben“ nicht einfach einen Ort, sondern sie bekommen einen Ort, wenn dieser Ort durch eine Messung erschaffen wird. Solange nicht der Ort eines Elektrons „etwa in der Mitte des Kristalls“ oder „ganz knapp links von der Plus-Elektrode“ oder wie auch immer durch eine Messung erschaffen wird, haben die Elektronen keinen Ort, d. h. sie sind (mindestens!) über den gesamten Kristall plus die beiden Elektroden delokalisiert.

Ein Modell mit delokalisierten Elektronen käme der Wahrheit näher; es wäre aber fürchterlich schwierig, mit dem verbesserten Modell den p-n-Übergang zu erklären. Das fehlerhafte Modell von Abb. 10.1 und 10.2 wird sich als gut geeignet zur Erklärung des p-n-Übergangs erweisen. Das liegt daran, dass es nicht nur falsche Züge (die Lokalisierung der Elektronen) sondern auch richtige Züge (das aus dem Naturgesetz (10.1b) folgende Verbot, einen Platz mit mehr als einem Elektron zu belegen) aufweist. Deshalb werden wir dies Modell trotz aller Bedenken doch weiterhin verwenden – mit der Vorsicht und Behutsamkeit, die bei einer **heuristischen** Argumentation geboten sind.⁸⁸

Heuristische Überlegungen können dann zum richtigen Ergebnis führen, wenn falsche Annahmen (hier: die Lokalisierung der Elektronen) durch geeignete weitere Annahmen kompensiert werden. Die beiden zusätzlichen Annahmen, mit denen das heuristische

⁸⁸ Die Autoren von Lehrbüchern für Elektronik-Ingenieure versuchen häufig, die Unzulänglichkeiten des Modells zu übertünchen, mithilfe kunstreicher Kombinationen von delokalisierten und lokalisierten Elektronen, sowie Energiebändern die durch Raumladungen verbogen werden. Ich halte es für besser, den Lesern reinen Wein einzuschenken und ohne viel Herumgerede den heuristischen Charakter dieses Halbleitermodells zuzugeben.

Modell von Abb. 10.1 und 10.2 zu korrekten Ergebnissen führt, sind die folgenden:

- * Metallische Elektroden können Elektronen in n-Halbleiter einspeisen, aber nicht in p-Halbleiter und nicht in undotierte Halbleiter.
- * Metallische Elektroden können Löcher in p-Halbleiter einspeisen, aber nicht in n-Halbleiter und nicht in undotierte Halbleiter.

Mit diesen beiden zusätzlichen Annahmen im Hinterkopf betrachten wir jetzt den in Abb. 10.3 auf der nächsten Seite skizzierten p-n-Übergang. Es handelt sich um einen Halbleiter-Kristall, dessen rechter Teil p-dotiert ist, und dessen linker Teil n-dotiert ist. Die n- und p-dotierten Bereiche grenzen unmittelbar aneinander.

In der oberen Skizze wird an die linke Elektrode eine höhere Spannung angelegt als an die rechte Elektrode. Folglich fließen die beweglichen Elektronen des n-Bereichs zur linken Elektrode ab, und die Löcher des p-Bereichs fließen zur rechten Elektrode ab. Wie in den beiden Zusatzannahmen formuliert, können aber weder von der rechten Elektrode Elektronen in den p-Halbleiter eingespeist werden, noch Löcher von der linken Elektrode in den n-Halbleiter. Also können die abgeflossenen Elektronen und Löcher nicht von außen ersetzt werden, und der Stromfluss kommt zum Erliegen. Durch das + und – beim Übergang wird angedeutet, dass sich im linken Teil des p-n-Übergangs (auf der n-Seite) eine positive Raumladung aufbaut (Mangel an Elektronen), und dass sich im rechten Teil des p-n-Übergangs (auf der p-Seite) eine negative Raumladung aufbaut (Mangel an Löchern).

In der unteren Skizze wurden die Elektroden umgepolt. Jetzt fließen die beweglichen Elektronen von links nach rechts in den p-Bereich, und die abgeflossenen Elektronen werden von der linken Elektrode ersetzt. Die Löcher fließen von rechts nach links in den

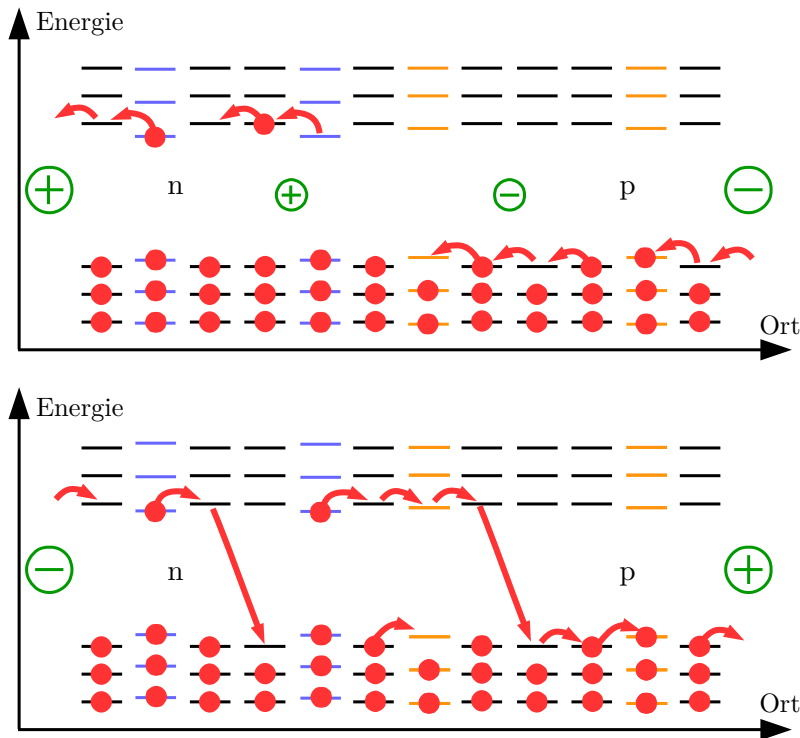


Fig. 10.3: p-n-Kontakt in Sperrrichtung (oben) und in Durchlassrichtung (unten)

n-Bereich, und die abgeflossenen Löcher werden von der rechten Elektrode ersetzt. Durch thermische Relaxation (Verlust von Energie durch Erzeugung von Vibrationen des Kristallgitters) werden die Elektronen früher oder später „in die Löcher herunterfallen“. Das stört den Stromtransport nicht, weil ständig neue Elektronen und Löcher von den Elektroden nachgeliefert werden.

Der p-n-Übergang wirkt also als ein Ventil für elektrische Ladung: Wenn der Pluspol am p-Halbleiter und der Minuspol am

n-Halbleiter liegt, dann wird Ladung durchgelassen, es fließt ein Strom. Bei umgekehrter Polung sperrt der p-n-Übergang, es fließt kein Strom. In der Elektrotechnik wird der p-n-Übergang als *Diode* bezeichnet.

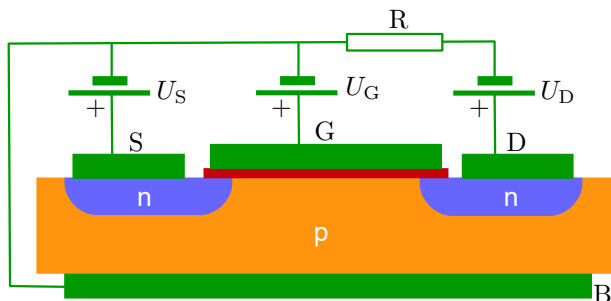


Fig. 10.4: Schnitt durch einen n-Kanal MOS-FET

In Abb. 10.4 ist ein Schnitt durch einen n-Kanal MOSFET (das **Akronym** steht für metal-oxide-semiconductor field-effect-transistor) gezeigt. In den gelb gemalten p-Halbleiter sind zwei blau gemalte Zonen mit n-Dotierung eindiffundiert. Es gibt vier grün gemalte Elektroden mit den Namen S = source, G = gate, D = drain, B = bulk.

Würde man an S eine negative und an B eine positive Spannung anlegen, dann hätte man eine Diode in Durchlassrichtung. Ebenso, wenn man an D eine negative und an B eine positive Spannung anlegen würde. Das tut man aber nicht, sondern man wählt die Source-Spannung U_S , die Drain-Spannung U_D , und die Gate-Spannung U_G stets

$$U_S \geq 0 \quad , \quad U_D \geq 0 \quad , \quad U_G \geq 0 .$$

Die Schaltzeichen oberhalb von Source, Gate, und Drain symbolisieren Spannungsquellen. Die Pluszeichen erinnern daran, dass U_S , U_D , und U_G niemals negativ eingestellt werden. (Sie können aber

Null sein.) Das Schaltzeichen R symbolisiert einen Widerstand, der dafür sorgt dass der Transistor nicht überlastet wird.

Wozu sind die beiden Dioden gut, wenn sie stets gesperrt sind? Nun, der Clou bei der Sache ist die Gate-Elektrode. Sie ist durch eine dünne, dunkelrot gemalte Isolationsschicht, die in der Regel aus SiO_2 = Siliziumdioxid besteht, vom Halbleiter getrennt. Dieser Aufbau des Gates aus den Schichten M = metal, O = oxide, S = semiconductor hat dem MOS-FET seinen Namen gegeben. Es gibt auch andere Arten von Feldeffekt-Transistoren, deren Gate anders konstruiert ist.

Solange U_G gleich Null ist fließt kein Strom, egal welche (positiven) Spannungen an Source und Drain gelegt sind. Denn die beiden Dioden sind ja stets gesperrt, es kann kein Strom durch die p-n-Übergänge fließen. Wenn aber $U_G > 0$ eingestellt wird, dann passiert etwas bemerkenswertes:

Gleichnamige Ladungen stoßen sich ab, ungleichnamige Ladungen ziehen sich an. Das positiv geladene Gate stößt die Löcher des p-Halbleiters ab, und zieht die Elektronen an. Das führt dazu, dass in einer hauchdünnen Schicht unterhalb des Isolators ein Zustand eintritt, der in Abb. 10.5 auf der nächsten Seite dargestellt wird: Im p-Bereich sind sämtliche Löcher verschwunden. Sämtliche Plätze unterhalb der Energielücke sind von Elektronen belegt.

Diese Graphik entspricht nahezu der unteren Graphik von Abb. 10.2 auf Seite 291. Es gibt nur zwei kleine Unterschiede: Erstens sind die Atomkerne der Al-Atome nur 13-fach positiv geladen, es befinden sich aber bei jedem Al-Atom von Abb. 10.3 14 Elektronen. Der p-Bereich enthält also eine kleine negative Raumladung, die jedoch den Stromtransport zwischen Source und Drain nicht ernsthaft behindert. Und zweitens gibt es im p-Bereich keine P-Atome. Das stört den Stromtransport noch weniger.

Wenn an Source und Drain unterschiedliche Spannungen liegen, dann können Elektronen sich wie in einem reinen n-Halbleiter auf

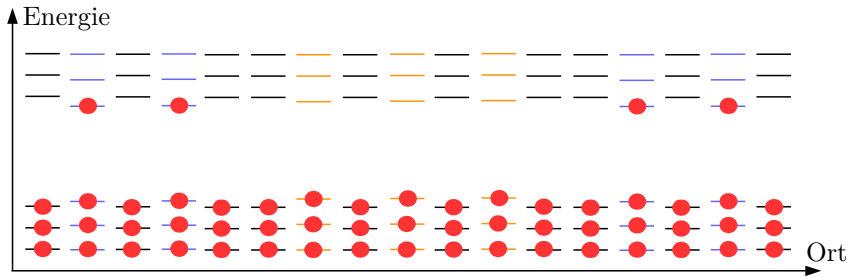


Fig. 10.5: Energieniveaus im n-Kanal

der einen Seite des Kristalls auf den Weg machen, und durch die freien Plätze oberhalb der Energielücke den langen Weg bis zur anderen Seite des Kristalls hinüber diffundieren. Es gibt keine Löcher, die die Elektronen einfangen könnten, denn sämtliche Plätze unterhalb der Energielücke sind bereits belegt.

Weil die Leitung des Stroms in der dünnen Schicht unter dem Isolator nahezu identisch mit der Stromleitung in einem reinen n-Halbleiter ist, wird diese Schicht als n-Kanal bezeichnet. Je nach Gatespannung ist der n-Kanal zwischen wenigen Nanometern und vielen hundert Nanometern dick. (In Leistungstransistoren kann der Kanal noch weitaus dicker sein.) Bei festgehaltener Differenz zwischen U_S und U_D wird der Strom zwischen Source und Drain um so größer, je dicker der n-Kanal ist. Durch das Gate fließt kein Strom, deshalb kann man durch Variation von U_G die Stromstärke zwischen Source und Drain nahezu verlustfrei regeln.⁸⁹ Ohne den Widerstand R würde der Transistor schnell Rauchzeichen geben und innerhalb weniger Sekunden zerstört sein. Man muss bei der Verwendung von Transistoren stets auf eine geeignet dimensionierte

⁸⁹ für Physiker: Die Regelung ist nur „nahezu“ verlustfrei, weil durch das Gate zwar kein Gleichstrom, wohl aber ein Wechselstrom fließt. In Hochfrequenz-Anwendungen sind die Schaltverluste des Gates ganz erheblich.

Beschaltung achten.

In jedem Prozessor eines Computers oder Smartphones sind viele Millionen von Feldeffekttransistoren zusammengeschaltet. Wären Elektronen nicht Fermionen (die der Fermi-Dirac-Statistik, sprich dem Naturgesetz (10.1b), unterliegen) sondern Dinge (die der Maxwell-Boltzmann-Statistik unterliegen), dann gäbe es keinen p-n-Übergang, keine Transistoren, keine Halbleiterelektronik.

10.3 Der Laser

Das Naturgesetz (10.1) hat zur Folge, dass beispielsweise zwei Elektronen (Elektronen sind Fermionen) in einem Festkörper niemals in allen Eigenschaften übereinstimmen können. Im Fall von Bosonen gibt es diese Einschränkung nicht. Das Licht eines Lasers besteht aus Photonen, Photonen sind Bosonen, und das Funktionsprinzip des Lasers besteht tatsächlich darin, dass eine gigantische Zahl von Photonen alle in genau dem gleichen Quantenzustand präpariert werden.

LASER ist ein **Akronym**: Es steht für Light Amplification by Stimulated Emission of Radiation. Um zu verstehen, was es mit der stimulierten Emission auf sich hat, müssen wir uns zunächst an einige wichtige Ergebnisse von Kapitel 3 erinnern:

Zur Deutung von Lenard's in Abschnitt 3.1 geschilderten Experimenten stellte Einstein 1905 seine Lichtquanten-Hypothese auf, d. h. er postulierte dass Atome und Moleküle Licht in Form unteilbarer Energiekörner absorbieren oder emittieren. Darüber habe ich in Abschnitt 3.2 berichtet. Einstein bezeichnete die Energiekörner als Lichtquanten. Später bürgerte sich dafür der Name Photonen ein.

Auf Seite 50 wurde die Gleichung (3.3) angegeben, mit der Planck im Jahr 1900 erstmals die Energiedichte der „Schwarzen Strahlung“ korrekt beschreiben konnte:

$$\text{Energiedichte der Schwarzen Strahlung} \stackrel{(3.3)}{=} \frac{8\pi h\nu^3/c^3}{e^{\frac{h\nu}{kT}} - 1} \quad (10.8)$$

Was mit „Schwarzer Strahlung“ gemeint ist, wurde in den Absätzen vor (3.3) erklärt. Diese Gleichung besagt, dass die Energiedichte der Strahlung mit der Frequenz ν in einem Ofen nur von der Temperatur T des Ofens abhängt. Alle anderen Faktoren in (10.8) sind Konstanten.

Bei der Veröffentlichung der Lichtquantenhypothese hatte Einstein darauf hingewiesen, dass diese Hypothese und Planck's Gleichung genau dann miteinander verträglich sind, wenn jedes Photon die Energie $h\nu$ hat, wobei ν die Frequenz des Lichts und h die Planck'sche Konstante ist. „Energiedichte“ bedeutet „Energie pro Volumen“. Man kann Planck's Strahlungsformel also auch als

$$\frac{N_{\text{Phot}} h\nu}{V} \stackrel{(10.8)}{=} \frac{8\pi h\nu^3/c^3}{e^{\frac{h\nu}{kT}} - 1} \quad (10.9)$$

schreiben, mit N_{Phot} gleich Anzahl der Photonen mit der Energie $h\nu$, die sich im Volumen V des Ofens befinden.

Im Jahr 1916 beschäftigte Einstein sich wieder einmal mit dieser merkwürdigen Formel. Dabei fiel ihm Folgendes auf [74]⁹⁰:

In Abbildung 10.6 auf der nächsten Seite wird der Vorgang von Absorption und Emission eines Photons symbolisch dargestellt. Die blaue Wellenlinie links symbolisiert ein Photon, das auf ein Atom in der Wand des Ofens trifft. Vereinfachend nehmen wir an, dass das Atom nur zwei unterschiedliche Zustände annehmen kann, in denen es die Energie E_j bzw. E_k hat. Die beiden möglichen Energien werden durch die schwarzen Striche dargestellt. Wenn das Atom die Energie E_j hat, ist der schwarze Strich j mit einem roten Punkt markiert; wenn es die Energie E_k hat, ist der Strich k markiert.

⁹⁰ Physiker finden in [75] eine elementare Darstellung von Einstein's Überlegungen.

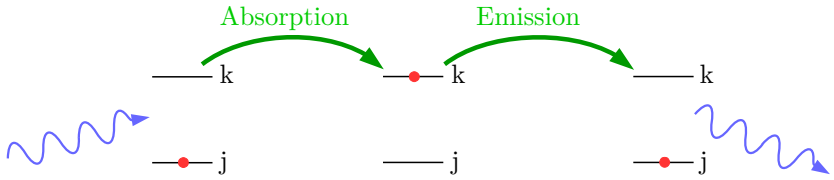


Abb. 10.6: Absorption und Emission eines Photons

Zwei Bedingungen müssen erfüllt sein, damit das Atom das Photon absorbieren kann: Erstens muss das Atom im Zustand j sein. Wenn es bei Ankunft des Photons bereits im Zustand k ist kann es kein Photon absorbieren, weil es in unserem einfachen Modell keinen möglichen Zustand mit noch größerer Energie als E_k gibt. Zweitens muss das Photon die passende Energie $h\nu = E_k - E_j$ haben.

Wenn beide Bedingungen erfüllt sind, dann *kann* das Atom das Photon absorbieren. Es wird das aber nicht mit Sicherheit tun, sondern nur mit einer gewissen Wahrscheinlichkeit, die von der Art des Atoms abhängt. Die Wahrscheinlichkeit W_{jk} dafür, dass in einer Sekunde irgendeines der Atome in den Wänden des Ofens ein Photon absorbiert ist **proportional** zur Zahl N_j dieser Atome, die sich gerade im Zustand j befinden, und proportional zur Dichte N_{Phot}/V der Photonen mit der passenden Energie $h\nu = E_k - E_j$ im Ofen. Die **Proportionalitätskonstante** nannte Einstein B_{jk} :

$$W_{jk} = B_{jk} N_j \frac{N_{\text{Phot}}}{V} \quad (10.10a)$$

Wenn das Atom das Photon absorbiert hat, dann hat es die Energie E_k . Nach einiger Zeit wird es ein Photon mit der Energie $h\nu = E_k - E_j$ emittieren, und anschließend nur noch die Energie E_j haben. Die Wahrscheinlichkeit W_{kj} dafür, dass in einer Sekunde irgendeines der Atome in der Wand des Ofens ein Photon emittiert,

ist proportional zur Zahl N_k der Atome, die sich gerade im Zustand k befinden. Die Proportionalitätskonstante, die ebenfalls von der Art der Atome abhängt, nannte Einstein A_{kj} :

$$W_{kj} = A_{kj} N_k \quad (10.10b)$$

Eine wichtige Eigenschaft des sogenannten Absorptionskoeffizienten B_{jk} und des Emissionskoeffizienten A_{kj} ist, dass beide nicht von der Temperatur abhängen. Das war bereits damals aus spektroskopischen Untersuchungen zuverlässig bekannt, und ist für Einstein's Schlussfolgerungen bedeutsam.

Wenn mehr Photonen von den Wänden des Ofens absorbiert als emittiert werden, dann nimmt die Zahl N_{Phot} der Photonen im Ofen ab. Umgekehrt nimmt die Zahl N_{Phot} der Photonen im Ofen zu, wenn die Wände mehr Photonen emittieren als absorbieren. Planck's Gleichung (10.9) beschreibt einen als „thermodynamisches Gleichgewicht“ bezeichneten Zustand, bei dem sich die Zahl der Photonen (bei konstanter Temperatur T) nicht verändert, d. h. es werden ständig gleich viel Photonen absorbiert und emittiert.

thermodynamisches Gleichgewicht:

$$B_{jk} N_j \frac{N_{\text{Phot}}}{V} \stackrel{(10.10a)}{=} W_{jk} = W_{kj} \stackrel{(10.10b)}{=} A_{kj} N_k \quad (10.11)$$

Aus dieser Gleichung folgt

thermodynamisches Gleichgewicht:

$$\frac{N_k}{N_j} \stackrel{(10.11)}{=} \frac{B_{jk} N_{\text{Phot}}/V}{A_{kj}} \quad (10.12)$$

Es gibt für das Verhältnis von N_k zu N_j im thermodynamischen Gleichgewicht eine weitere Relation, die damals ebenfalls bereits wohlbekannt und fest etabliert war:

thermodynamisches Gleichgewicht:

$$\frac{N_k}{N_j} = e^{-(E_k - E_j)/(kT)} \begin{cases} < 1 & \text{bei } T < \infty \\ \rightarrow 1 & \text{bei } T \rightarrow \infty \end{cases} \quad (10.13)$$

Leser, die die Exponentialfunktion nicht kennen, sollten einfach glauben dass die Temperaturabhängigkeit so ist, wie hier angegeben: Bei jeder endlichen Temperatur ist im thermodynamischen Gleichgewicht $N_k/N_j < 1$, und das Verhältnis N_k/N_j kommt der 1 um so näher, je höher die Temperatur ist.

Die Anzahl der Photonen im Ofen steigt nach Planck's Gleichung mit zunehmender Temperatur T immer weiter an, und erreicht bei $T \rightarrow \infty$ beliebig große Werte:

$$\frac{N_{\text{Phot}}}{V} \stackrel{(10.9)}{=} \frac{1}{h\nu} \cdot \frac{8\pi h\nu^3/c^3}{e^{\frac{h\nu}{kT}} - 1} \xrightarrow{T \rightarrow \infty} \infty \quad (10.14)$$

(Wieder sollten Leser, die die Exponentialfunktion $e^{\frac{h\nu}{kT}}$ nicht kennen, das einfach glauben.) Mit diesem Ergebnis folgt aus (10.12)

thermodynamisches Gleichgewicht:

$$\frac{N_k}{N_j} \stackrel{(10.12)}{=} \frac{B_{jk} N_{\text{Phot}}/V}{A_{kj}} \xrightarrow{T \rightarrow \infty} \infty. \quad (10.15)$$

Hier stimmt etwas nicht! Laut (10.13) kann das Verhältnis N_k/N_j im thermodynamischen Gleichgewicht niemals größer als 1 sein, aber laut (10.15) wird es bei hoher Temperatur beliebig groß.

Wo steckt der Fehler? Einstein wollte – zu Recht, wie wir heute wissen – weder an der Relation (10.13) noch an Planck's Gleichung (10.14) zweifeln. Bleiben als mögliche Fehlerquellen nur noch die beiden Gleichungen (10.10a) und (10.10b). Einstein tippte auf (10.10b), und zeigte dass alle Unstimmigkeiten verschwinden wenn man diese Gleichung durch

$$W_{kj} = A_{kj}N_k + B_{kj}N_k \frac{N_{\text{Phot}}}{V} \quad (10.16)$$

ersetzt. Der Term $A_{kj}N_k$ war bereits in (10.10b) enthalten. Er wird als „spontane Emission“ bezeichnet, und ist im rechten Teil von Abb. 10.6 auf Seite 302 symbolisiert. Neu hinzugekommen ist der Term $B_{kj}N_k N_{\text{Phot}}/V$, der als „stimulierte Emission“ bezeichnet und in Abb. 10.7 veranschaulicht wird.

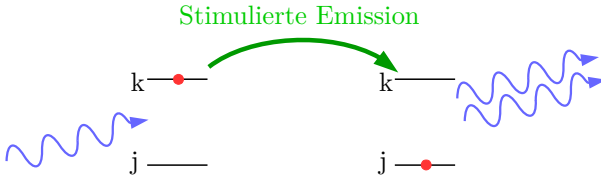


Abb. 10.7: Stimulierte Emission eines Photons

Bei der stimulierten Emission trifft ein Photon auf ein Atom, das sich bereits im angeregten Zustand k befindet. Also kann das Atom das Photon nicht absorbieren. Oben haben wir angenommen dass das Atom das Photon einfach ignoriert, aber Einstein postulierte dass das nicht stimmt. Vielmehr spürt das angeregte Atom die Anwesenheit des Photons, und wird dadurch zur Emission eines weiteren Photons stimuliert. Die Wahrscheinlichkeit W_{kj} dafür, dass das in einer Sekunde irgendwo an den Wänden des Ofens geschieht, ist proportional zur Anzahl N_k der Atome, die sich gerade im Zustand k befinden, und proportional zur Dichte N_{Phot}/V der Photonen im Ofen, die die Energie $h\nu = E_k - E_j$ haben. Die Proportionalitätskonstante nannte Einstein B_{kj} .

Mit (10.16) erhält man im thermodynamischen Gleichgewicht anstelle von (10.11)

thermodynamisches Gleichgewicht:

$$\begin{aligned}
 B_{jk}N_j \frac{N_{\text{Phot}}}{V} &\stackrel{(10.10a)}{=} W_{jk} = W_{kj} \stackrel{(10.16)}{=} A_{kj}N_k + B_{kj}N_k \frac{N_{\text{Phot}}}{V} \\
 \implies \frac{N_k}{N_j} &= \frac{B_{jk}}{A_{kj}/(N_{\text{Phot}}/V) + B_{kj}} \quad (10.17)
 \end{aligned}$$

Jeder der drei Koeffizienten A_{kj} , B_{jk} , und B_{kj} ist größer als Null, und von der Temperatur unabhängig. Die Photonendichte N_{Phot}/V ist ebenfalls größer als Null, und steigt bei $T \rightarrow \infty$ gegen unendlich an. Also hat (10.17) die richtige, mit (10.13) übereinstimmende Temperaturabhängigkeit, wenn die Koeffizienten für Absorption und für stimulierte Emission gleich sind:

$$B_{jk} = B_{kj} \quad (10.18)$$

Einstein konnte zeigen⁹¹ dass man von der gut bekannten Eigenschaft (10.13) des thermodynamischen Gleichgewichts direkt zur Struktur (10.9) der Planck'schen Gleichung gelangt, wenn man annimmt dass es stimulierte Emission mit der Eigenschaft (10.18) gibt. Das war ein bedeutender Fortschritt, weil dies die erste plausible Erklärung dieser Gleichung war, die Planck im Jahr 1900 als „glücklich erratene Interpolationsformel“ gefunden hatte.

Einstein gab sich mit diesem wissenschaftlichen Erfolg zufrieden. Eine technische Nutzung der stimulierten Emission zog damals niemand in Betracht. Das geschah erst seit den fünfziger Jahren. 1953 realisierten Charles Townes (1915–2015) und James Gordon (1928–2013) einen Laser, der jedoch keinen Lichtstrahl sondern einen Strahl von Mikrowellen erzeugte, und deshalb als Maser

⁹¹ für Physiker: $\frac{N_k}{N_j} \stackrel{(10.13)}{=} e^{-(E_k - E_j)/(kT)} \stackrel{(10.17)}{=} \frac{B_{jk}}{A_{kj}/(N_{\text{Phot}}/V) + B_{kj}}$
 $\implies \frac{N_{\text{Phot}}}{V} = \frac{A_{kj}/B_{jk}}{e^{+(E_k - E_j)/(kT)} - B_{kj}/B_{jk}}$

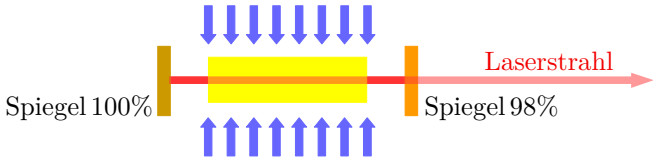


Abb. 10.8: Das Funktionsprinzip eines Lasers

bezeichnet wurde. Den ersten sichtbaren Laserstrahl erzeugte im Jahr 1960 Theodore Maiman (1927–2007).

In Abb. 10.8 ist skizziert, wie ein Laser im Prinzip funktioniert. Zwischen zwei Spiegeln befindet sich das (hier gelb angedeutete) Laser-Medium. Das Medium kann ein Gas sein, oder eine Flüssigkeit, oder ein Festkörper. Durch die blauen Pfeile wird die Anregung des Mediums angedeutet, die zum Beispiel durch externe Beleuchtung oder durch eine elektrische Gasentladung realisiert werden kann.

Was dann geschieht, kann man aus Gleichung (10.17) ablesen:

$$B_{jk}N_j \frac{N_{\text{Phot}}}{V} = W_{jk} \stackrel{(10.17)}{=} W_{kj} = A_{kj}N_k + B_{kj}N_k \frac{N_{\text{Phot}}}{V}$$

Die angeregten Atome des Mediums emittieren spontan — proportional zum Faktor A_{kj} — Photonen. Die meisten Photonen fliegen in irgend eine Richtung davon und gehen verloren. Aber bald wird eines der Photonen zufällig so auf die Spiegel treffen, dass es ins Medium zurückgespiegelt wird, und vielfach zwischen den beiden Spiegeln hin und her reflektiert wird. Dies Photon wird dann — proportional zum Faktor $B_{kj}N_k$ — die Emission eines weiteren Photons stimulieren. Oder es wird — proportional zum Faktor $B_{jk}N_j$ — von einem Atom des Mediums absorbiert werden.

Was man gerne möchte ist, dass das zwischen den Spiegeln hin und her fliegende Photon möglichst viele angeregte Atome zur Emission eines weiteren Photons stimuliert. Diese sollen dann ebenfalls zwischen den Spiegeln hin und her fliegen und weitere angeregte Atome zur Emission weiterer Photonen stimulieren, so dass man

insgesamt eine lawinenartige Verstärkung des Photonenstrahls zwischen den Spiegeln erhält. Das wird aber im thermodynamischen Gleichgewicht nicht geschehen, denn dann ist ja laut (10.13) stets $N_k/N_j \leq 1$. Folglich ist im

thermodynamischen Gleichgewicht:

$$\frac{\text{Wahrscheinlichkeit für stimulierte Emission}}{\text{Wahrscheinlichkeit für Absorption}} =$$

$$\stackrel{(10.17)}{=} \frac{B_{kj}N_k}{B_{jk}N_j} \stackrel{(10.18)}{=} \frac{N_k}{N_j} \stackrel{(10.13)}{<} 1 \quad (10.19)$$

Das zwischen den Spiegeln hin und her fliegende Photon wird mit höherer Wahrscheinlichkeit gleich wieder absorbiert, als dass es eine stimulierte Emission auslöst. Eine Verstärkung des Lichtstrahls zwischen den Spiegeln kann es nur mit $N_k > N_j$ geben, d. h. wenn *kein* thermodynamisches Gleichgewicht besteht.

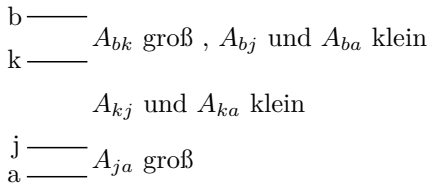


Abb. 10.9: Ein 4-Niveau Lasermedium

Man kann sich von der störenden Einschränkung durch das thermodynamische Gleichgewicht befreien, wenn man zur Anregung des Mediums durch die äußere Lichtquelle oder die elektronische Entladung andere Energieniveaus wählt als für die stimulierte Emission. In Abb. 10.9 wird das Prinzip eines 4-Niveau-Lasers gezeigt. Man regt möglichst viele Atome aus dem Grundzustand a in den Zustand b an. Wenn die Koeffizienten für spontane Emission so sind wie in Abb. 10.9 eingetragen, dann werden zwar wie beim 2-Niveau-System stets mehr Atome im Zustand a als im Zustand b

sein ($N_a > N_b$), aber zugleich werden mehr Atome im Zustand k als im Zustand j sein ($N_j < N_k$). Denn der Zustand k ist langlebig (wegen A_{kj} und A_{ka} klein), und wird (wegen A_{bk} groß, A_{bj} klein, A_{ba} klein) ständig aus dem Zustand b effizient bevölkert. Dagegen ist der Zustand j kurzlebig (wegen A_{ja} groß), und wird ständig effizient in den Zustand a entleert.

Dank der sogenannten Besetzungsinversion $N_j < N_k$ bewirken Photonen mit der Energie $h\nu = E_k - E_j$ häufiger stimulierte Emission, als dass sie selbst absorbiert werden:

$$\frac{\text{Wahrscheinlichkeit für stimulierte Emission}}{\text{Wahrscheinlichkeit für Absorption}} =$$

$$\stackrel{(10.17)}{=} \frac{B_{kj}N_k}{B_{jk}N_j} \stackrel{(10.18)}{=} \frac{N_k}{N_j} > 1 \text{ im System Abb. 10.9} \quad (10.20)$$

Also wird man eine Verstärkung des Laserlichts mit der Photonenenergie $h\nu = E_k - E_j$ erhalten.

Das durch Stimulation emittierte Photon hat den gleichen Quantenzustand wie das Photon, das die stimulierte Emission bewirkt hat. An dieser Stelle ist wichtig, dass Photonen als Bosonen dem Naturgesetz (10.1a) unterliegen. Wenn Photonen Fermionen wären, dann könnten laut (10.1b) niemals zwei von ihnen den gleichen Quantenzustand haben. Die beiden Photonen haben den gleichen Quantenzustand, bewegen sich also im Takt, mit Wellenberg bei Wellenberg und Wellental bei Wellental, und in die gleiche Richtung, wie rechts in Abb. 10.7 auf Seite 305 angedeutet.

Bei jedem Durchgang durch das Lasermedium stimulieren diese Photonen die Emission weiterer Photonen, so dass nach kurzer Zeit der Raum zwischen den beiden Spiegeln mit einer gigantischen Menge von Photonen gefüllt ist, die (fast) alle im Takt schwingen, und (fast) alle die gleiche Energie $h\nu = E_k - E_j$ haben. Ein kleiner Teil dieser Photonen (z. B. 2% im Beispiel von Abb. 10.8) wird durch einen der beiden Spiegel ausgekoppelt, und bildet den nutzbaren Laserstrahl.

Weil die Wellenberge und Wellentäler der Photonen im Laserstrahl aneinander ausgerichtet sind, wird Laserlicht in populären Darstellungen zuweilen als „kohärent“ bezeichnet, im Gegensatz zum Licht beispielsweise der Sonne oder von Glühlampen, das „inkohärent“ sei. In dieser Sprechweise wird der Sachverhalt bis ins Unsinnige vereinfacht.

Es ist nicht so, dass manche Sorten von Licht „kohärent“ und andere Sorten „inkohärent“ sind. Sondern jedes Licht hat eine bestimmte *Kohärenzlänge*, die man mit einem Interferometer, wie es in Abb. 2.2 auf Seite 24 dargestellt wurde, ausmessen kann. Dazu justiert man den beweglichen Schlitten S zunächst so ein, dass die Wege A und B genau gleich lang sind. Dann beginnt man den Schlitten zu verschieben. Dabei beobachtet man dass die Lichtintensität, die von den Detektoren D_G und D_H gemessen wird, zunächst variiert wie in Abb. 2.3 dargestellt. Aber bei größerer Weglängendifferenz werden die Interferenzen immer undeutlicher, und schließlich beobachten beide Detektoren nur noch gleichmäßig die halbe Gesamtintensität, unabhängig von der weiteren Verschiebung des Schlittens.

Die maximale Weglängendifferenz, bei der die Interferenzen noch klar erkennbar⁹² sind, ist die Kohärenzlänge des Lichts. Jedes Licht, auch Laserlicht, hat nur eine endliche Kohärenzlänge. Das liegt daran, dass im Lasermedium die spontane Emission zwar wesentlich seltener vorkommt als die stimulierte Emission, aber nicht völlig verschwunden ist. Und jedes Licht, auch das Licht der Sonne und das Licht von Glühlampen, hat⁹³ eine Kohärenzlänge > 0 . Das liegt daran dass auch ein Photon, das sich nur einmal durch das strahlende Medium bewegt, eine kleine Chance hat stimulierte Emission bei anderen angeregten Atomen zu bewirken.

⁹² Man muss natürlich ein präzises quantitatives Kriterium dafür definieren, was mit „klarer Erkennbarkeit“ genau gemeint ist.

⁹³ Mit Licht der Kohärenzlänge = 0 würde kein optisches Instrument funktionieren, nicht einmal eine einfache Lesebrille.

Wenn man bei der Messung der Kohärenzlänge von Sonnenlicht als Detektor einfach sein menschliches Auge verwendet, dann findet man als Kohärenzlänge etwa $1 \mu\text{m}$. Wenn man dagegen Detektoren verwendet, die auch infrarotes und ultraviolettes Licht wahrnehmen, dann wird die Kohärenzlänge kleiner. Größer wird die Kohärenzlänge, wenn man vor den Detektor einen Farbfilter hält, der z. B. nur grünes Licht oder nur rotes Licht durchlässt.

Es gibt demnach einen eindeutigen Zusammenhang zwischen der Bandbreite des (wahrgenommenen) Lichts und seiner Kohärenzlänge: Je monochromatischer das Licht ist (d. h. je kleiner seine wahrgenommene Bandbreite ist), desto größer ist die Kohärenzlänge. Wenn die Bandbreite des Detektors größer ist als die Bandbreite des untersuchten Lichts, dann wird die Kohärenzlänge allein durch die Bandbreite des untersuchten Lichts bestimmt. Das ist z. B. der Fall, wenn man mit dem menschlichen Auge als Detektor die Kohärenzlänge des Lichts ausmisst, das Gase bei niedrigem Druck in elektrischen Entladungen emittieren.

Den „Weltrekord“ der Kohärenzlängen hielt vor der Entwicklung der Laser das rote Licht der Cäsium-Gasentladung mit etwa 2 m. Mit stabilen Lasern kann man Kohärenzlängen von vielen Dutzenden von Kilometern erreichen. Das zeigt dass Laser unglaublich schmalbandiges (monochromatisches) Licht emittieren können.

Man könnte schmalbandiges Licht auch dadurch herstellen, dass man breitbandiges Licht durch einen Filter schickt, der nur Licht mit einer sehr schmalen Bandbreite durchlässt. Aber dies Licht wäre ziemlich schwach, weil ja der größte Teil der Intensität im Filter absorbiert wird. Bei Laserlicht braucht man keine Kompromisse zu schließen: Man kann ohne Filterverluste sehr, sehr schmalbandiges Licht mit sehr, sehr hoher Intensität bekommen. Viele Anwendungen der Lasertechnik in Wissenschaft, Kommunikationstechnik, und Vermessungstechnik beruhen auf dieser einzigartigen Kombination von schmaler Bandbreite und hoher Intensität.

This page is intentionally empty.

Encyclopedia

5^{-3} : a^b is read as “a to the power of b”. The number in the superscript is the **exponent**. The exponent indicates how many times the number below should be multiplied by itself:



$$3^4 = 3 \cdot 3 \cdot 3 \cdot 3 = 81$$

$$17 \cdot 10^7 = 17 \cdot 10\,000\,000 = 170\,000\,000$$

If the exponent has a negative sign, the reciprocal is meant:

$$5^{-3} = \frac{1}{5^3} = \frac{1}{5 \cdot 5 \cdot 5} = \frac{1}{125}$$

$$67\,258 \cdot 10^{-2} = \frac{67\,258}{10^2} = \frac{67\,258}{100} = 672.58$$

back by the   key combination, or by the back-button of the pdf-reader



Acronym: (Greek) akron = peak, summit, (Greek) onoma = name
 An acronym is an abbreviation formed from the initial characters of several words. Examples:

NYSE = New York Stock Exchange

PC = personal computer

UK = United Kingdom

SPDC = spontaneous parametric down conversion

back by the   key combination, or by the back-button of the pdf-reader

$\alpha, \beta, \gamma, \dots$ Greek characters:

α = alpha

ι = iota

Π = Pi

β = beta

κ = kappa

ρ = rho

γ = gamma

λ = lambda

σ = sigma

Γ = Gamma

Λ = Lambda

Σ = Sigma

δ = delta

μ = mu

τ = tau

Δ = Delta

ν = nu

ϕ, φ = phi

ϵ, ε = epsilon

ξ = xi

χ = chi

ζ = zeta

Ξ = Xi

ψ = psi

η = eta



o = omikron

ω = omega

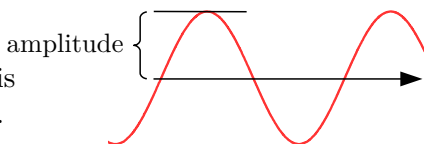
θ, ϑ = theta



π = pi

Ω = Omega

back by the   key combination, or by the back-button of the pdf-reader

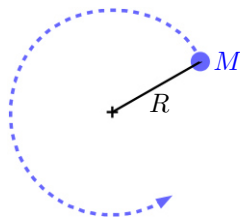
Amplitude: The amplitude is the maximum value of a wave.



back by the   key combination, or by the back-button of the pdf-reader

Angular Momentum: If an object of mass M moves at a speed of v in a circular path with radius R around the center, then its angular momentum \mathbf{J} is equal to



$$|\mathbf{J}| = R \cdot \text{momentum} = R \cdot M \cdot v .$$



Like [momentum](#), angular momentum has a magnitude and a direction in space, and is therefore represented in theory by a vector and denoted in bold. The direction of angular momentum is *not* defined as the direction of momentum (since that changes permanently along the circular path), but rather as the direction of the axis of rotation. Specifically, this is analogous to the direction of motion of a right-handed screw: If I turn the screw in the direction of the dashed line in the sketch, it comes out of the board and moves toward me. Accordingly, in the example of the sketch, \mathbf{J} is perpendicular to the plane of the paper and directed toward the observer. If, on the other hand, I turn the screw clockwise, it moves away from me and into the board. Accordingly, \mathbf{J} is directed perpendicular to the plane of the paper away from the observer when M moves clockwise.



Angular momentum is a conserved quantity, just like momentum. Many people are familiar with this from merry-go-rounds: if you

walk from the outside toward the center (toward the axis of rotation) on the rotating platform, the merry-go-round spins faster. If you walk away from the axis of rotation toward the outside, the carousel spins more slowly. As the distance R decreases, the speed v must increase, and vice versa, so that the angular momentum $|\mathbf{J}| = R \cdot M \cdot v$ remains constant.



back by the   key combination, or by the back-button of the pdf-reader

anisotropic: (Greek) iso = equal, (Greek) tropos = direction, an- = (negation).



Isotropic means, that all directions in space are equal or equivalent. Anisotropic means the opposite.

back by the   key combination, or by the back-button of the pdf-reader

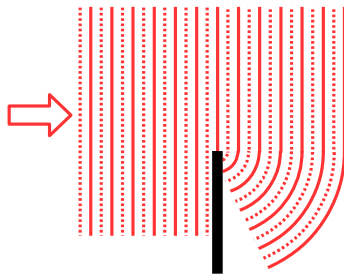
to compensate: (Latin) compensare = to balance something



back by the   key combination, or by the back-button of the pdf-reader

credo: (Latin) credo = I believe



back by the   key combination, or by the back-button of the pdf-reader

Diffraction: Diffraction refers to the bending of waves around an obstacle. In the diagram, the solid lines represent wave crests, and the dashed lines represent wave troughs. The wave propagates from left to right. Behind the black obstacle, it spreads due to *diffraction* into the region of space that would in a straight line be blocked by the obstacle. Diffraction occurs due to constructive or destructive interference of partial waves that pass at different distance above the obstacle.



back by the   key combination, or by the back-button of the pdf-reader



elastically: A collision is called “elastic” if no kinetic energy is converted into heat. The opposite is an “inelastic” collision, in which — example: colliding automobiles — a greater or lesser portion of the kinetic energy is converted into heat, or in which — example: colliding atoms — a greater or lesser portion of the kinetic energy is converted into “internal degrees of freedom” (excitation of electrons).

back by the   key combination, or by the back-button of the pdf-reader

Electrode: (Greek) odos = path, street



→ electrode = “street for electricity”

Thus electrodes are wires or strips that conduct electricity well. Usually they are made of metal.

back by the   key combination, or by the back-button of the pdf-reader



eV: 1 eV = 1 electron-Volt is the kinetic energy of an electron, which has been accelerated by a voltage of 1 V = 1 Volt, resp. which can be brought to rest by a reverse voltage of 1 V.

$$1 \text{ eV} = 6.62606876 \cdot 10^{-34} \text{ Joule} = 6.62606876 \cdot 10^{-34} \frac{\text{kg m}^2}{\text{s}^2}$$



back by the   key combination, or by the back-button of the pdf-reader

empirical: (Greek:) empeiria = experience



→ empirical = based on experience, derived from experience

back by the   key combination, or by the back-button of the pdf-reader

ex cathedra: (Latin) ex cathedra = from the chair (of a professor). A cathedral is a church building where a bishop proclaims eternal truths from his chair. When something is proclaimed “ex cathedra”, it is best to keep any doubts to oneself; otherwise, one could quickly find oneself at the stake.



back by the   key combination, or by the back-button of the pdf-reader

Frequency: Frequency is defined as the number of oscillations (or other processes that repeat at regular intervals) per unit of time. The standard unit is: Hz = Hertz = per second = s^{-1}



back by the   key combination, or by the back-button of the pdf-reader

heuristic:



(Greek) ἐυρίσκω (spoken: heurisko) = I discover, I detect
A Method for the solution of a scientific or mathematical problem is called heuristic, if it ultimately leads to the correct solution despite following dubious or even flawed paths.

back by the   key combination, or by the back-button of the pdf-reader



Hexagon: (Greek) hex = six ; (Greek) gonia = corner, angle

back by the   key combination, or by the back-button of the pdf-reader

Ion: An electron has an electric charge of -1.6×10^{-19} Coulomb, and a proton has an electric charge of $+1.6 \times 10^{-19}$ Coulomb. If the number of electrons in an atom's shell is equal to the number of protons in its nucleus (which is the normal case), then the atom is electrically neutral overall. If $n = 1, 2, 3, \dots$ electrons are removed from an atom, then its charge is $+n \cdot 1 \times 10^{-19} \text{ C}$, and it is referred to as an n -fold positively charged ion. If n electrons are added to an atom, its charge is $-n \cdot 1.6 \cdot 10^{-19} \text{ C}$, and it is referred to as an n -fold negatively charged ion. As free particles, negatively charged ions are very rare because they are extremely unstable. As components of chemical compounds, however, they occur frequently.



back by the   key combination, or by the back-button of the pdf-reader

Isotope: The nuclei of atoms are composed of protons and neutrons. Each proton has an electric charge of $+1.6 \cdot 10^{-19}$ Coulomb. Neutrons are electrically neutral. Many elements occur with different numbers of neutrons. For example, the atomic nucleus of chlorine always has 17 protons, but sometimes 18 and sometimes 20 neutrons. To distinguish between these two *isotopes*, the total number of protons and neutrons (i. e. 35 or 37 in the case of chlorine) is written as a superscript before the chemical symbol of the element, i. e. ^{35}Cl or ^{37}Cl in the case of chlorine.



back by the   key combination, or by the back-button of the pdf-reader

Mathematical Symbols:

$=$ equal	$>$ greater than
\approx approx. equal	\geq equal or greater than
\sim proportional to	\gtrsim approx. equal or slightly greater than
$\hat{=}$ corresponds to	$<$ less than
∞ infinite	\leq equal or less than
	\lesssim approx. equal or slightly less than



back by the   key combination, or by the back-button of the pdf-reader

Measuring Eyepiece: In a microscope (and also in a telescope), the eyepiece refers to the lens — or, in modern instruments, the lens system — that faces the observer's eye, as opposed to the objective lens, which faces the object being observed. A measuring eyepiece is an eyepiece into which a scale has been engraved or etched, allowing the size of the observed object to be measured.

back by the   key combination, or by the back-button of the pdf-reader

micro-,kilo-,Giga-,... Prefixes for powers of ten:

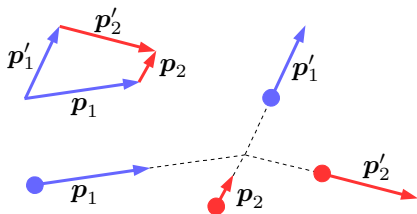
	symbol	factor	example
nano-	n	10^{-9}	5.4 nm = 5.4 nanometer = 0.000 000 005 4 m
micro-	μ	10^{-6}	2.7 μ s = 2.7 microseconds = 0.000 002 7 s
milli-	m	10^{-3}	3 mm = 3 millimeter = 0.003 m
centi-	c	10^{-2}	9.1 cm = 9.1 centimeter = 0.091 m
dezi-	d	10^{-1}	2 dl = 2 deziliter = 0.2 l
hecto-	h	10^2	1013 hPa = 1013 hectopascal = 101 300 Pa
kilo-	k	10^3	4.1 kg = 4.1 kilogram = 4 100 g
Mega-	M	10^6	1.3 MW = 1.3 Megawatt = 1 300 000 W
Giga-	G	10^9	12 GHz = 12 Gigahertz = 12 000 000 000 Hz

back by the   key combination, or by the back-button of the pdf-reader

Momentum: If an object with mass M is moving at speed v , this is it's

$$\text{momentum : } \quad \mathbf{p} = M \cdot \mathbf{v}$$

The momentum is a conserved quantity. If the blue and the red particle with momenta \mathbf{p}_1 and \mathbf{p}_2 collide, then for their momenta \mathbf{p}'_1 and \mathbf{p}'_2 after the collision holds due to





$$\text{momentum conservation:} \\ \mathbf{p}'_1 + \mathbf{p}'_2 = \mathbf{p}_1 + \mathbf{p}_2$$

Regarding momentum conservation, not only the magnitude (represented by the length of the arrows) of the momenta but also their direction must be considered; i. e. the arrows must be added geometrically, as shown in the sketch top left. To emphasize this, the momenta \mathbf{p} are printed bold.

The mass M of the objects is the relativistic mass, which is connected with the mass M_0 of the objects at rest due to

$$M = \frac{M_0}{\sqrt{1 - v^2/c^2}} .$$

v^2 is the square of the velocity of the object, and c^2 is the square of the speed of light in vacuum. If $v < 0.1 \cdot c \approx 3 \cdot 10^7 \text{m/s}$, the difference in-between M and M_0 is tiny, and can be neglected in most cases.

back by the   key combination, or by the back-button of the pdf-reader

nm, ns:

1 nm = 1 nanometer = 10^{-9} m = 0.000 000 001 m

1 ns = 1 nanosecond = 10^{-9} s = 0.000 000 001 s

back by the   key combination, or by the back-button of the pdf-reader

orthogonal: (Greek) orthos = right, straight ;
(Greek) gonia = corner, angle

back by the   key combination, or by the back-button of the pdf-reader



Pentagon:

(Greek) pente = five ; (Greek) gonia = corner, angle

back by the   key combination, or by the back-button of the pdf-reader

piezoelectric actuator: (Greek) piezo = I press, I squeeze,
(Latin) actio = action



When pressure is applied to certain crystals or ceramics, a positive electric charge appears on one surface and an equal negative charge on the opposite surface. Conversely, when an external electric voltage is applied to the surfaces of these solids, they contract or expand by up to about one-tenth of a percent of their thickness, depending on the magnitude of the applied voltage. This effect is utilized in piezoelectric actuators.

back by the   key combination, or by the back-button of the pdf-reader

Probability: The probability of an event that is certain to occur is 1. The probability of an event that is impossible to occur is 0. The probability of an event that may or may not occur is greater than zero and less than one.

Defining the concept of “probability” without falling into circular reasoning is one of the most difficult problems there is. Philosophers, mathematicians, and scientists have written thick books on the subject, yet still haven’t been able to reach a complete consensus.

Sometimes debonair naivety is an advantage, and in this case it certainly is. We will use in this book the term “probability” just as any reasonably intelligent housewife would: we will not define it at all, because “we already know what it means”. This pragmatic approach will serve us perfectly well.

back by the   key combination, or by the back-button of the pdf-reader

proportional: A quantity A is proportional to a quantity B (formal notation: $A \sim B$), if

$$A = f \cdot B \quad \text{with } f = \text{constant} .$$


The factor f is called proportionality constant.

Example: Let a stone fall down near Earth surface. It's velocity is proportional to the time elapsed since it started to fall:

$$\text{velocity} = g \cdot \text{time elapsed since start of fall}$$

$$\text{with } g = 9.81 \frac{\text{m}}{\text{s}^2} = \text{constant}$$

$$\text{velocity} \sim \text{time elapsed since start of fall}$$

back by the   key combination, or by the back-button of the pdf-reader

rational, irrational: (Latin) ratio = mind, reason; or:

or: ratio, quotient. irratio = the opposite of ratio



The word rational (resp. irrational) is in use with both these different meanings.

ratio = mind, reason: A person's behavior or arguments are rational (and thus predictable) if they follow reasonable (i. e., understandable) rules. Otherwise, they are irrational. A natural process is rational if it proceeds in accordance with laws of nature that can, in principle, be discovered, so that the process can be calculated. If the process is not governed by any law of nature, then it is, in principle, unpredictable and is called irrational.

ratio = ratio, quotient: A number r is called rational, if it is the quotient



$$r = \frac{m}{n} \quad \text{with } m, n = 0, \pm 1, \pm 2, \pm 3, \dots$$

of two integers m and n . Numbers, which are not the quotient of two integers, are called irrational. Example: The square root of 2 is an irrational number.

back by the   key combination, or by the back-button of the pdf-reader



reflect: (Latin) re = back, (Latin) flectere = to bow, to bend

The prefix “re-” is traditionally interpreted not so strictly in physics. Even if the light is only deflected to the side, this is still referred to as reflection.



back by the   key combination, or by the back-button of the pdf-reader

rhetorical: (Latin) rhetor = the speaker, the orator



When speakers ask rhetorical questions, they don't want to hear answers, because they are about to provide them by themselves. The questions are just a trick to capture the audience's attention.

back by the   key combination, or by the back-button of the pdf-reader



to scatter: The verb "to scatter" is used when radiation or a projectile is deflected in a different direction by an obstacle. For example, in a game of billiards, one ball is scattered by another billiard ball.

back by the   key combination, or by the back-button of the pdf-reader


Schizophrenia: (Greek) schizein = to split, (Greek) phren = mind
In modern psychiatry, schizophrenia is diagnosed (in my humble opinion) when an obvious mental disorder defies all established classifications and cannot be categorized under any other diagnosis.

back by the   key combination, or by the back-button of the pdf-reader

Single-Crystal: In a single-crystal solid (also called mono-crystal solid), the regular arrangement of atoms extends throughout the entire volume, whereas a polycrystalline solid is composed of a usually very large number of crystallites. Within a single crystallite, the atoms are arranged regularly, but the various crystallites within the polycrystalline solid are randomly twisted and offset relative to one another.



back by the   key combination, or by the back-button of the pdf-reader

synchronous: (Greek) syn- = together-, (Greek) chronos = time, → synchronous = at the same time, simultaneous



back by the   key combination, or by the back-button of the pdf-reader

Syntax and semantics: (Greek) sema = symbol, character
(Greek) syn- = together-, (Greek) taxis = order

Syntax is the study of the correct arrangement of symbols; semantics is the study of the meaning of symbols.

back by the   key combination, or by the back-button of the pdf-reader

Synthesis: (Greek) syn- = together-, (Greek) thesis = position
→ synthesis = composition

back by the   key combination, or by the back-button of the pdf-reader

Appendix

A.1 Are photons divisible wave packets?

If light were composed of divisible wave packets, then (3.14) had to be ≥ 1 ,

$$(3.14) = \frac{N_{\text{GTR}}N_{\text{G}}}{N_{\text{GT}}N_{\text{GR}}} \geq 1$$

for the following reason. The Cauchy-Schwarz inequality⁹⁴, which we take without explanation from the mathematical formula collection, states that the mean of the square of any quantity I_{E} is always greater than or equal to the square of its mean value:

$$\overline{I_{\text{E}} \cdot I_{\text{E}}} \geq \overline{I_{\text{E}}} \cdot \overline{I_{\text{E}}} .$$

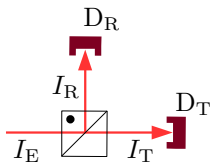
This can also be written as

$$\frac{\overline{I_{\text{E}} \cdot I_{\text{E}}}}{\overline{I_{\text{E}}} \cdot \overline{I_{\text{E}}}} \geq 1 . \quad (\text{A.1})$$

The $\overline{\quad}$ is to symbolize the mean value.

Now we tentatively assume that light consists of divisible wave packets. We will call the intensity (as a reminder: intensity = power is the energy carried by light per unit time) of the light entering the beam splitter I_{E} . At the beam splitter, half of the light is transmitted and half is reflected:

⁹⁴ This inequality has been proven by Augustin-Louis Cauchy (1789–1857) and Hermann Amandeus Schwarz (1843–1921).



$$I_T = \frac{1}{2} I_E \quad (\text{A.2a})$$

$$I_R = \frac{1}{2} I_E \quad (\text{A.2b})$$

$$I_T \cdot I_R = \frac{1}{4} I_E \cdot I_E \quad (\text{A.2c})$$

I_T is the intensity of the transmitted light, and I_R is the intensity of the reflected light. Thereby the Cauchy-Schwarz inequality (A.1) can be written as follows:

$$1 \leq \frac{\overline{I_E \cdot I_E}}{\overline{I_E} \cdot \overline{I_E}} = \frac{4 \overline{I_T \cdot I_R}}{2 \overline{I_T} \cdot 2 \overline{I_R}} \quad (\text{A.3})$$

If energy is conserved (i. e. if no energy can appear out of nowhere and no energy can disappear to nowhere, which no physicist would doubt without compelling reason), then the probability W_T that within a 2.5 ns time window detector D_T will trigger, must be proportional to the average intensity of the light reaching this detector during this time window:

$$W_T \sim \overline{I_T} \quad (\text{A.4a})$$

Accordingly the probability W_R that detector D_R triggers within the same time interval, and the probability W_{TR} that both detectors trigger within the same time interval, must be

$$W_R \sim \overline{I_R} \quad (\text{A.4b})$$

$$W_{TR} \sim \overline{I_T \cdot I_R} \quad (\text{A.4c})$$

This is inserted into (A.3):

$$1 \leq \frac{\overline{I_T \cdot I_R}}{\overline{I_T} \cdot \overline{I_R}} = \frac{W_{TR}}{W_T \cdot W_R} \quad (\text{A.5})$$

W_T is equal to the number N_T of events counted by detector D_T , divided by the number N_G of all observed 2.5 ns long time windows. The same holds for W_R and W_{TR} :

$$W_T = N_T/N_G \quad (\text{A.6a})$$

$$W_R = N_R/N_G \quad (\text{A.6b})$$

$$W_{TR} = N_{TR}/N_G \quad (\text{A.6c})$$

Thereby (A.5) becomes

$$1 \leq \frac{W_{TR}}{W_T \cdot W_R} = \frac{N_{TR}/N_G}{(N_T/N_G) \cdot (N_R/N_G)} = (\text{3.14}) . \quad (\text{A.7})$$

(3.14) had to be ≥ 1 , if light was composed of divisible wave-packets. The measured result, however, was

$$(\text{3.14}) = 0.0177 \pm 0.0026 .$$

By this measured result, the hypothesis of divisible wave packets is definitively disproved.

→ back to page 64

A.2 Drehimpuls und Polarisation von Photonen

Drehimpuls ist eine Erhaltungsgröße, d. h. Drehimpuls kann weder aus dem Nichts auftauchen noch im Nichts verschwinden. Wenn sich durch Absorption oder Emission eines Photons der Drehimpuls eines Atoms ändert, dann muss das Photon diesen Drehimpuls mitgebracht bzw. mitgenommen haben. Aus dem in Abb. A.1 auf der nächsten Seite nochmal abgedruckten Termschema des Calcium-Atoms erkennt man, dass sich bei der Emission der beiden Lumineszenz-Photonen die Drehimpuls-Quantenzahl j des Atoms jeweils um ± 1 ändert. Diese beiden Photonen sind deshalb zirkular polarisiert.

Was zirkuläre Polarisation ist wird in Abb. A.2 auf Seite 334 erklärt. In dieser Graphik sind verschiedene Wellen gezeichnet. In

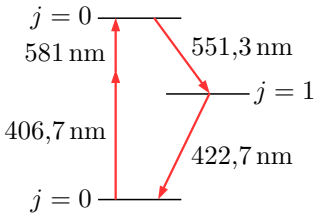


Abb. A.1: Eine Lumineszenz-Kaskade des Calcium-Atoms

jedem Diagramm ist die Summe der gestrichelt gezeichneten und der gepunktet gezeichneten Welle als dicke durchgezogene Linie dargestellt. In den Diagrammen auf der rechten Seite sieht man die gleichen Wellen, aber jetzt mit Blickrichtung parallel zur z -Achse.

Die gestrichelte und die gepunktete Welle schwingen stets in einer Richtung: Die gestrichelte Welle parallel zur y -Achse, die gepunktete Welle parallel zur x -Achse. Diese Art der Polarisation wird als „linear“ bezeichnet und mit dem Buchstaben L gekennzeichnet. Die gestrichelte Welle ist L_0 -polarisiert, denn der Winkel zwischen ihrer Schwingungsrichtung und der y -Achse ist 0° . Die gepunktete Welle ist L_{90} -polarisiert, denn der Winkel zwischen ihrer Schwingungsrichtung und der y -Achse ist 90° . In den beiden oberen Diagrammen auf der rechten Seite von Abb. A.2 erkennt man, wie der Winkel zwischen der y -Achse und der Schwingungsrichtung von Wellen definiert ist, nämlich gemessen von der positiven y -Achse gegen den Uhrzeigersinn.

In Diagramm A.2(a) schwingen die gestrichelte L_0 -Welle und die gepunktete L_{90} -Welle „in Phase“. Das bedeutet, dass beide Wellen ihre Maxima, Minima, und Nulldurchgänge jeweils an den gleichen Stellen der z -Achse haben. Ihre Summe, die als durchgezogene Linie gezeichnete L_{45} -Welle, ist ebenfalls linear polarisiert.

In Diagramm A.2(b) läuft die gepunktete Welle der gestrichelten um eine halbe Wellenlänge nach, falls sich die Welle in positiver z -Richtung ausbreitet (bei $z = 0$ hat die L_0 -Welle ihr Maximum bereits eine viertel Wellenlänge hinter sich, während die L_{90} -Welle

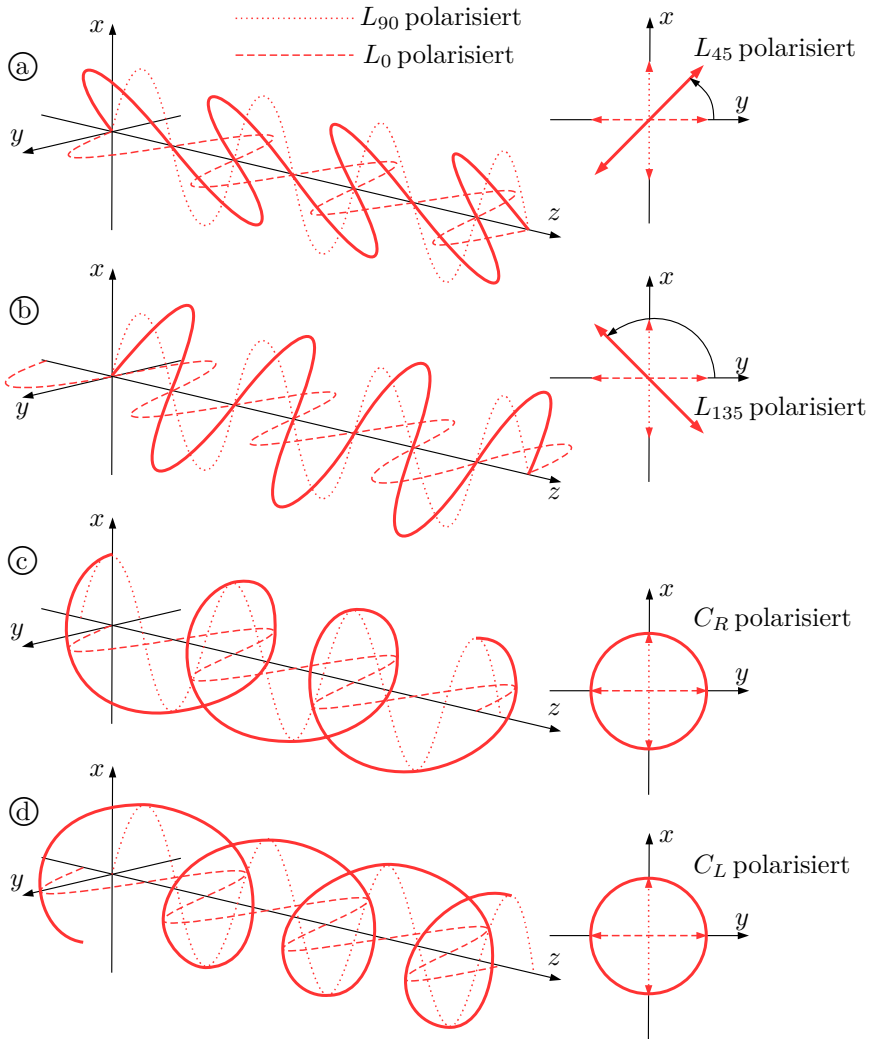


Fig. A.2: Linear und zirkular polarisierte Wellen

ihr Maximum noch eine viertel Wellenlänge vor sich hat), bzw. um eine halbe Wellenlänge voraus, falls sich die Welle in negativer z -Richtung ausbreitet (bei $z = 0$ hat die L_{90} -Welle ihr Maximum bereits eine viertel Wellenlänge hinter sich, während die L_0 -Welle ihr Maximum noch eine viertel Wellenlänge vor sich hat). Auch in diesem Fall ist die Summe der beiden Teilwellen eine linear polarisierte Welle, aber jetzt eine Welle mit L_{135} -Polarisation.

In Diagramm A.2(c) läuft die gepunktete Welle der gestrichelten um eine viertel Wellenlänge voraus, falls sich die Welle in positiver z -Richtung ausbreitet (bei $z = 0$ hat die L_{90} -Welle ihr Maximum bereits erreicht, während die L_0 -Welle ihr Maximum noch eine viertel Wellenlänge vor sich hat), bzw. um eine viertel Wellenlänge hinterher, falls sich die Welle in negativer z -Richtung ausbreitet (bei $z = 0$ hat die L_0 -Welle ihr Maximum bereits eine viertel Wellenlänge hinter sich, während die L_{90} -Welle ihr Maximum gerade erst erreicht). In diesem Fall bildet die Summe der beiden Teilwellen eine Rechtsschraube. Man⁹⁵ bezeichnet diese Welle als „rechts-zirkular“ polarisiert bzw. C_R -polarisiert.⁹⁶

In Diagramm A.2(d) schließlich ist eine links-zirkular polarisierte (C_L -polarisierte) Welle dargestellt. Sie entsteht, wenn die L_{90} -

⁹⁵ „Man“ ist jeder Mensch, der schon mal eine Schraube in ein Brett geschraubt hat. Diese Erfahrung kann man bei Theoretischen Physikern nicht unbedingt voraussetzen. Bei der theoretischen Analyse des Vorgangs haben einige Theoretiker sich überlegt, dass man ja auch die Schraube still halten und stattdessen das Brett drehen kann, um die Schraube einzudrehen. Folglich bezeichnen sie eine Schraube als Rechtsschraube, wenn man das Brett rechts herum drehen muss, damit es die Schraube in sich einsaugt. Und sie bezeichnen eine Schraube als Linksschraube, wenn man das Brett links herum drehen muss, damit es die Schraube in sich einsaugt. Konsequenterweise verwenden sie dann auch die Bezeichnungen rechts-zirkular und links-zirkular umgekehrt wie wir. Hinweis: Wenn ein Theoretiker noch niemals eine Schraube in ein Brett geschraubt hat beweist das noch lange nicht, dass sein Optik-Lehrbuch schlecht ist.

⁹⁶ Der Buchstabe C steht für (lateinisch) circus = Kreis.

polarisierte Teilwelle der L_0 -polarisierten Teilwelle um eine viertel Wellenlänge hinterher läuft falls die Welle sich in positiver z -Richtung ausbreitet (bei $z = 0$ hat die L_{90} -Welle ihr Maximum bereits eine viertel Wellenlänge hinter sich, während die L_0 -Welle ihr Maximum gerade erst erreicht), bzw. um eine viertel Wellenlänge voraus läuft falls sich die Welle in negativer z -Richtung ausbreitet (bei $z = 0$ hat die L_{90} -Welle ihr Maximum bereits eine viertel Wellenlänge hinter sich, während die L_0 -Welle ihr Maximum gerade erst erreicht).

Wie ein Phasenversatz in Bildern zu malen ist, sieht man in Abb. A.2. Aber wie wird ein Phasenversatz mathematisch in unsere Zustandsvektoren eingefügt? Da verkünde ich einfach mal **ex cathedra** folgende Regel:⁹⁷

$$\begin{aligned}
 \text{Phasenversatz } + 1/4 \text{ Wellenlänge} &\longleftrightarrow \text{Faktor } + \sqrt{-1} = +i \\
 \text{Phasenversatz } - 1/4 \text{ Wellenlänge} &\longleftrightarrow \text{Faktor } - \sqrt{-1} = -i \\
 \text{Phasenversatz } \pm 1/2 \text{ Wellenlänge} &\longleftrightarrow \text{Faktor } - 1 \\
 \text{Phasenversatz } 0 &\longleftrightarrow \text{Faktor } + 1
 \end{aligned} \tag{A.8}$$

Hier wird die Zahl $i = \sqrt{-1}$ verwendet, die in (5.5) erklärt wurde.

Also kann man den Zustandsvektor $|L_{45}\rangle$ eines L_{45} -polarisierten Photons als Summe eines L_0 -polarisierten Photons und eines L_{90} -polarisierten Photons mit Phasenversatz Null folgendermaßen schreiben:

$$|L_{45}\rangle = \sqrt{\frac{1}{2}} \left(|L_0\rangle + |L_{90}\rangle \right) \tag{A.9a}$$

Der Faktor $\sqrt{1/2}$ ist erforderlich, damit die Projektions-Amplitude dieses Zustandsvektors auf sich selbst 1 ist, wie es sein muss. Prüfen wir es nach:

⁹⁷ für Physiker: Die allgemeine Regel ist: Phasenversatz $\varphi \longleftrightarrow$ Faktor $e^{i\varphi}$

$$\begin{aligned} \langle L_{45} || L_{45} \rangle &= \\ &= \frac{1}{2} \left(\underbrace{\langle L_0 || L_0 \rangle}_1 + \underbrace{\langle L_0 || L_{90} \rangle}_0 + \underbrace{\langle L_{90} || L_0 \rangle}_0 + \underbrace{\langle L_{90} || L_{90} \rangle}_1 \right) = 1 \end{aligned}$$

Mit den Phasenversatz-Faktoren (A.8) sind die Zustandsvektoren von Photonen mit L_{135} -, C_R -, und C_L -Polarisation:

$$|L_{135}\rangle = \sqrt{\frac{1}{2}} \left(|L_0\rangle - |L_{90}\rangle \right) \quad (\text{A.9b})$$

$$|C_R\rangle = \sqrt{\frac{1}{2}} \left(|L_0\rangle + i|L_{90}\rangle \right) \text{ falls das} \quad (\text{A.9c})$$

Photon sich in $+z$ -Richtung bewegt

$$|C_R\rangle = \sqrt{\frac{1}{2}} \left(|L_0\rangle - i|L_{90}\rangle \right) \text{ falls das} \quad (\text{A.9d})$$

Photon sich in $-z$ -Richtung bewegt

$$|C_L\rangle = \sqrt{\frac{1}{2}} \left(|L_0\rangle - i|L_{90}\rangle \right) \text{ falls das} \quad (\text{A.9e})$$

Photon sich in $+z$ -Richtung bewegt

$$|C_L\rangle = \sqrt{\frac{1}{2}} \left(|L_0\rangle + i|L_{90}\rangle \right) \text{ falls das} \quad (\text{A.9f})$$

Photon sich in $-z$ -Richtung bewegt

Wer als Übungsaufgabe nachprüfen will, ob auch die Projektionsamplituden dieser Zustandsvektoren auf sich selbst gleich 1 sind, muss beachten dass im linken Faktor von Projektionsamplituden das Vorzeichen von i umgedreht wird, und dass $(+i) \cdot (-i) \stackrel{(5.5g)}{=} +1$ ist:

$$\langle C_R || C_R \rangle \stackrel{(\text{A.9c})}{=} \sqrt{\frac{1}{2}} \left(\langle L_0 | - i \langle L_{90} | \right) \sqrt{\frac{1}{2}} \left(|L_0\rangle + i|L_{90}\rangle \right)$$

$$\langle C_R || C_R \rangle \stackrel{(\text{A.9d})}{=} \sqrt{\frac{1}{2}} \left(\langle L_0 | + i \langle L_{90} | \right) \sqrt{\frac{1}{2}} \left(|L_0\rangle - i|L_{90}\rangle \right)$$

$$\begin{aligned} \langle C_L || C_L \rangle &\stackrel{\text{(A.9e)}}{=} \sqrt{\frac{1}{2}} \left(\langle L_0 | + i \langle L_{90} | \right) \sqrt{\frac{1}{2}} \left(|L_0\rangle - i |L_{90}\rangle \right) \\ \langle C_L || C_L \rangle &\stackrel{\text{(A.9f)}}{=} \sqrt{\frac{1}{2}} \left(\langle L_0 | - i \langle L_{90} | \right) \sqrt{\frac{1}{2}} \left(|L_0\rangle + i |L_{90}\rangle \right) \end{aligned}$$

Damit zurück zum Term-Schema von Caesium auf Seite 333. Beim Übergang vom obersten zum mittleren Zustand und beim Übergang vom mittleren zum unteren Zustand emittiert das Atom jeweils ein Photon, wobei sich seine Drehimpuls-Quantenzahl um ± 1 ändert. Man sieht es den zirkular polarisierten Wellen A.2 © und A.2 ④ schon intuitiv an, dass ihr Drehimpuls von Null verschieden ist, und dieser intuitive Eindruck ist auch völlig richtig. Ohne Beweis teile ich hier mit, dass die Emission eines zirkular polarisierten Photons tatsächlich die Drehimpuls-Quantenzahl j eines Atoms um ± 1 ändert. Die beiden Lumineszenz-Photonen Photon₁ und Photon₂ müssen also C_R - oder C_L -polarisiert sein. Die Kaskade der zwei Lumineszenz-Photonen startet und endet bei einem Zustand des Atoms mit Drehimpuls-Quantenzahl $j = 0$. Also müssen sich die Drehimpulse von Photon₁ und Photon₂ gerade kompensieren. Da liegt die Vermutung nahe, dass eines der Photonen C_R -polarisiert ist, und das andere C_L -polarisiert ist.

Aber Halt! Das gilt nur wenn die beiden Lumineszenz-Photonen sich in gleicher Richtung bewegen. Aspect et al. untersuchten jedoch den Fall, in dem die beiden Lumineszenz-Photonen in genau entgegengesetzten Richtungen emittiert werden, siehe Abb. 7.1 auf Seite 160. In diesem Fall müssen entweder beide Photonen rechtszirkular polarisiert oder beide linkszirkular polarisiert sein, damit sich ihre Drehimpulse zu Null addieren. Ein anschauliches Beispiel sind zwei Leute, die gleichzeitig von gegenüberliegenden Seiten jeder eine Schraube in das gleiche Brett drehen. Wenn der eine eine Rechtsschraube verwendet und der andere eine Linksschraube, dann müssen sie das Brett festhalten damit es sich nicht mitdreht. Aber wenn beide eine Rechtsschraube oder beide eine Linksschrau-

be verwenden und **synchron** mit gleicher Kraft schrauben, dann bleibt das Brett in Ruhe, auch wenn es nicht festgehalten wird. Ebenso ist die Summe der Drehimpulse von zwei zirkular polarisierten Photonen, die sich in entgegengesetzter Richtung bewegen, genau dann Null, wenn entweder beide C_R -polarisiert oder beide C_L -polarisiert sind.

Wegen dieser Korrelation (beide C_R -polarisiert oder beide C_L -polarisiert) wird das Gesamtsystem der zwei Lumineszenz-Photonen in der Quantentheorie durch den verschränkten Zustandsvektor

$$|\text{Photon}_1\&\text{Photon}_2\rangle = \sqrt{\frac{1}{2}} \left(|C_R\rangle_1 |C_R\rangle_2 + |C_L\rangle_1 |C_L\rangle_2 \right) \quad (\text{A.10a})$$

beschrieben. Hier wurden die Indizes von Photon_1 und Photon_2 an die Vektoren $|C_R\rangle$ und $|C_L\rangle$ angefügt, damit man erkennt welcher Vektor zu welchem Photon gehört.

In (A.9) wurden die Vektoren $|C_R\rangle$ und $|C_L\rangle$ als Kombinationen von $|L_0\rangle$ und $|L_{90}\rangle$ geschrieben. Das setzen wir in den verschränkten Zustandsvektor (A.10a) ein:

$$\begin{aligned} |\text{Photon}_1\&\text{Photon}_2\rangle &\stackrel{(\text{A.10a})}{=} \sqrt{\frac{1}{2}} \left[|C_R\rangle_1 |C_R\rangle_2 + |C_L\rangle_1 |C_L\rangle_2 \right] = \\ &= \sqrt{\frac{1}{2}} \left[\sqrt{\frac{1}{2}} \left(|L_0\rangle_1 + i|L_{90}\rangle_1 \right) \sqrt{\frac{1}{2}} \left(|L_0\rangle_2 - i|L_{90}\rangle_2 \right) + \right. \\ &\quad \left. + \sqrt{\frac{1}{2}} \left(|L_0\rangle_1 - i|L_{90}\rangle_1 \right) \sqrt{\frac{1}{2}} \left(|L_0\rangle_2 + i|L_{90}\rangle_2 \right) \right] = \\ &= \sqrt{\frac{1}{2}} \left[\frac{1}{2} \left(|L_0\rangle_1 |L_0\rangle_2 - i|L_0\rangle_1 |L_{90}\rangle_2 + \right. \right. \\ &\quad \left. \left. + i|L_{90}\rangle_1 |L_0\rangle_2 - i^2 |L_{90}\rangle_1 |L_{90}\rangle_2 \right) + \frac{1}{2} \left(|L_0\rangle_1 |L_0\rangle_2 + \right. \right. \\ &\quad \left. \left. + i|L_0\rangle_1 |L_{90}\rangle_2 - i|L_{90}\rangle_1 |L_0\rangle_2 - i^2 |L_{90}\rangle_1 |L_{90}\rangle_2 \right) \right] = \\ &= \sqrt{\frac{1}{2}} \left(|L_0\rangle_1 |L_0\rangle_2 + |L_{90}\rangle_1 |L_{90}\rangle_2 \right) \end{aligned} \quad (\text{A.10b})$$

Hier wurde $-i^2 \stackrel{(5.5g)}{=} +1$ benutzt. Jetzt müssen wir noch eine Verallgemeinerung einfügen: Aus Abb. A.2 erkennt man, dass wir genau die gleichen Wellen mit zirkularer Polarisation C_R und C_L als Kombination von zwei linear polarisierten Wellen mit Polarisation $L_{\gamma+0}$ und $L_{\gamma+90}$ erhalten könnten, mit beliebigem Winkel γ . Es kommt nur darauf an, dass der Winkel zwischen den beiden linear polarisierten Wellen genau 90° ist. Statt A.10b müssen wir also die allgemeine Lösung

$$|\text{Photon}_1 \& \text{Photon}_2\rangle = \sqrt{\frac{1}{2}} \left(|L_\gamma\rangle_1 |L_\gamma\rangle_2 + |L_{\gamma+90}\rangle_1 |L_{\gamma+90}\rangle_2 \right) \\ \text{mit beliebigem } \gamma \quad (\text{A.10c})$$

einsetzen.

Der Gesamt-Zustandsvektor von zwei Photonen, die in entgegengesetzte Richtung fliegen und entweder beide rechtszirkular oder beide linkszirkular polarisiert sind, ist also identisch mit dem Gesamt-Zustandsvektor von zwei Photonen, die in entgegengesetzte Richtung fliegen und entweder beide linear L_γ -polarisiert oder beide linear $L_{\gamma+90}$ -polarisiert sind, mit beliebigem Winkel γ bzw. $\gamma + 90^\circ$ zwischen der y-Achse und der Polarisationsebene der Photonen.⁹⁸ Außerdem müssen die Teilwellen $|L_\gamma\rangle$ und $|L_{\gamma+90}\rangle$ um genau eine viertel Wellenlänge gegeneinander versetzt sein, damit sie in der Summe eine zirkular polarisierte Welle ergeben, siehe Abb. A.2. Für die Auswertung des Experiments von Aspect et al.

⁹⁸ Es lohnt sich, darüber etwas genauer nachzudenken. Emittiert das Atom die beiden Photonen nun „in Wirklichkeit“ mit zirkularer oder mit linearer Polarisation? Die Antwort der Quantentheorie lautet: Weder noch. Die Polarisation der Photonen wird erst durch die Messung erschaffen. Wenn man Messgeräte verwendet die lineare Polarisation erschaffen, dann werden die Photonen nach der Messung linear polarisiert sein, und ihre Polarisation wird gemäß (A.10c) korreliert sein. Wenn man Messgeräte verwendet die zirkulare Polarisation erschaffen, dann werden die Photonen nach der Messung zirkular polarisiert sein, und ihre Polarisation wird gemäß (A.10a) korreliert sein.

genügt aber die Messung der linearen Polarisation der Photonen und ihrer Korrelation gemäß (A.10c), den Phasenversatz kann man ignorieren.

→ zurück zu Gleichung (7.1)

A.3 Berechnung der Wahrscheinlichkeiten (7.4)

Wir wollen die Wahrscheinlichkeiten

$$W_{RR} \stackrel{(7.4a)}{=} \frac{1}{2} \left| \langle a_1 || L_\gamma \rangle_1 \langle a_2 || L_\gamma \rangle_2 + \langle a_1 || L_{\gamma+90} \rangle_1 \langle a_2 || L_{\gamma+90} \rangle_2 \right|^2$$

$$W_{RT} \stackrel{(7.4b)}{=} \frac{1}{2} \left| \langle a_1 || L_\gamma \rangle_1 \langle b_2 || L_\gamma \rangle_2 + \langle a_1 || L_{\gamma+90} \rangle_1 \langle b_2 || L_{\gamma+90} \rangle_2 \right|^2$$

$$W_{TR} \stackrel{(7.4c)}{=} \frac{1}{2} \left| \langle b_1 || L_\gamma \rangle_1 \langle a_2 || L_\gamma \rangle_2 + \langle b_1 || L_{\gamma+90} \rangle_1 \langle a_2 || L_{\gamma+90} \rangle_2 \right|^2$$

$$W_{TT} \stackrel{(7.4d)}{=} \frac{1}{2} \left| \langle b_1 || L_\gamma \rangle_1 \langle b_2 || L_\gamma \rangle_2 + \langle b_1 || L_{\gamma+90} \rangle_1 \langle b_2 || L_{\gamma+90} \rangle_2 \right|^2$$

der vier möglichen Messergebnisse berechnen. Dazu stellen wir zunächst fest, dass

$$|L_\gamma\rangle = e^{i\gamma} |L_0\rangle \quad \text{und} \quad |L_{\gamma+90}\rangle = e^{i\gamma} |L_{90}\rangle \quad (\text{A.12})$$

ist. Ich weiß, dass die Exponentialfunktion mit imaginärem Exponenten hier ziemlich unvermittelt kommt. Nicht-Physiker mögen bitte einfach mal glauben, dass (A.12) tatsächlich korrekt ist. Außerdem können wir $|L_0\rangle$ bzw. $|L_{90}\rangle$ durch die Einheitsvektoren $|y\rangle$ bzw. $|x\rangle$ in Richtung der Koordinatenachsen y bzw. x ersetzen, denn so wurden die Polarisationsvektoren ja definiert, siehe Abb. 7.3 auf Seite 162. Die Indizes 1 und 2, mit denen die Zugehörigkeit der Funktionen zu Photon₁ und Photon₂ spezifiziert werden, sind nicht mehr erforderlich nachdem sie mit den jeweils passenden Gegenstücken zu Projektionsamplituden zusammengefasst wurden. Eine weitere Vereinfachung ergibt sich daraus, dass $|e^{i\gamma}|^2 = 1$ ist.

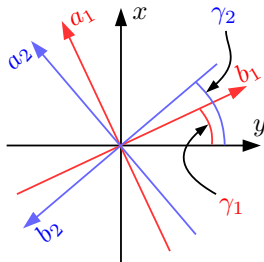


Fig. A.3: Die drei Koordinatensysteme

Also können die Wahrscheinlichkeiten folgendermaßen geschrieben werden:

$$W_{RR} = \frac{1}{2} \left| \langle a_1 || y \rangle \langle a_2 || y \rangle + \langle a_1 || x \rangle \langle a_2 || x \rangle \right|^2 \quad (\text{A.13a})$$

$$W_{RT} = \frac{1}{2} \left| \langle a_1 || y \rangle \langle b_2 || y \rangle + \langle a_1 || x \rangle \langle b_2 || x \rangle \right|^2 \quad (\text{A.13b})$$

$$W_{TR} = \frac{1}{2} \left| \langle b_1 || y \rangle \langle a_2 || y \rangle + \langle b_1 || x \rangle \langle a_2 || x \rangle \right|^2 \quad (\text{A.13c})$$

$$W_{TT} = \frac{1}{2} \left| \langle b_1 || y \rangle \langle b_2 || y \rangle + \langle b_1 || x \rangle \langle b_2 || x \rangle \right|^2 \quad (\text{A.13d})$$

Die Einheitsvektoren $|a_1\rangle$, $|b_1\rangle$, $|a_2\rangle$, $|b_2\rangle$, $|x\rangle$, $|y\rangle$ sind alle reell. In Gleichung (5.7) auf Seite 109 hatten wir festgestellt, dass für solche Vektoren gilt:

falls $\langle h || g \rangle$ reell ist:

$$\langle h || g \rangle \stackrel{(5.7)}{=} \langle g || h \rangle \stackrel{(5.7)}{=} \cos \angle(h, g) \quad (\text{A.14})$$

Damit wir von (A.14) Gebrauch machen können, sind in Abb. A.3 die drei Koordinatensysteme von Abb. 7.4 noch einmal gezeichnet, aber jetzt alle drei mit Blickrichtung parallel zur z -Achse, und durch unterschiedliche Farben deutlicher hervorgehoben. Mithilfe

dieser Grafik erkennt man:⁹⁹

$$\begin{aligned}
 \langle a_1 || y \rangle &= \cos(\gamma_1 + 90^\circ) \\
 \langle a_2 || y \rangle &= \cos(\gamma_2 + 90^\circ) \\
 \langle a_1 || x \rangle &= \langle b_1 || y \rangle = \cos(\gamma_1) \\
 \langle a_2 || x \rangle &= \cos(\gamma_2) \\
 \langle b_2 || y \rangle &= \cos(\gamma_2 + 180^\circ) \stackrel{\text{Abb. 5.4}}{=} -\cos(\gamma_2) \\
 \langle b_2 || x \rangle &= \cos(\gamma_2 + 90^\circ) \\
 \langle b_1 || x \rangle &= \cos(90^\circ - \gamma_1) \stackrel{\text{Abb. 5.4}}{=} -\cos(90^\circ + \gamma_1)
 \end{aligned}$$

Also kann man (A.13) folgendermaßen schreiben:

$$W_{RR} = \frac{1}{2} \left| \cos(\gamma_1 + 90^\circ) \cos(\gamma_2 + 90^\circ) + \cos(\gamma_1) \cos(\gamma_2) \right|^2 \quad (\text{A.16a})$$

$$W_{RT} = \frac{1}{2} \left| -\cos(\gamma_1 + 90^\circ) \cos(\gamma_2) + \cos(\gamma_1) \cos(\gamma_2 + 90^\circ) \right|^2$$

$$W_{TR} = \frac{1}{2} \left| \cos(\gamma_1) \cos(\gamma_2 + 90^\circ) - \cos(90^\circ + \gamma_1) \cos(\gamma_2) \right|^2$$

$$W_{TT} = \frac{1}{2} \left| -\cos(\gamma_1) \cos(\gamma_2) - \cos(90^\circ + \gamma_1) \cos(\gamma_2 + 90^\circ) \right|^2$$

Offensichtlich ist

⁹⁹ In (6.1) traten bei den Projektionsamplituden Winkel auf, die nur halb so groß waren wie die Winkel im Ortsraum. Das liegt daran dass man einen Stern-Gerlach-Magneten um 180° drehen muss damit sein Eigenvektor $|\uparrow\rangle$ in seinen Eigenvektor $|\downarrow\rangle$ übergeht, während der Winkel zwischen $|\uparrow\rangle$ und $|\downarrow\rangle$ im abstrakten mathematischen Raum der Eigenvektoren 90° beträgt. Dagegen muss man die Polarisatoren nur um 90° im Ortsraum drehen, damit ihre Eigenvektoren $|a\rangle$ in die Eigenvektoren $|b\rangle$ übergehen. Im abstrakten mathematischen Raum der Eigenvektoren ist der Winkel zwischen $|a\rangle$ und $|b\rangle$ ebenfalls 90° . In diesem Fall sind also die Winkel im Ortsraum und im abstrakten Vektorraum identisch.

$$W_{RT} = W_{TR} , \quad (\text{A.17a})$$

und wegen $|-1|^2 = +1$ auch

$$W_{TT} = W_{RR} . \quad (\text{A.17b})$$

Weil jedes gültige Ergebnis mit Sicherheit (also mit Wahrscheinlichkeit $W = 1$) eines der vier möglichen sein muss, gilt schließlich auch noch

$$1 = W_{RR} + W_{RT} + W_{TR} + W_{TT}$$

$$\stackrel{(\text{A.17a}), (\text{A.17b})}{=} 2W_{RR} + 2W_{RT}$$

$$\frac{1}{2} = W_{RR} + W_{RT}$$

$$W_{RT} = \frac{1}{2} - W_{RR} \quad (\text{A.17c})$$

$$W_{TR} \stackrel{(\text{A.17a})}{=} \frac{1}{2} - W_{RR} . \quad (\text{A.17d})$$

Man braucht also nur W_{RR} zu berechnen, dann kennt man mithilfe von (A.17) auch die anderen drei Wahrscheinlichkeiten.

Aus der Formelsammlung entnehmen wir die Formel:

$$\cos(\gamma_1 + 90^\circ) \cos(\gamma_2 + 90^\circ) + \cos(\gamma_1) \cos(\gamma_2) = \cos(\gamma_1 - \gamma_2) \quad (\text{A.18})$$

Damit erhält man:

$$W_{RR} \stackrel{(\text{A.16a})}{=} \frac{1}{2} \left| \cos(\gamma_1 + 90^\circ) \cos(\gamma_2 + 90^\circ) + \cos(\gamma_1) \cos(\gamma_2) \right|^2$$

$$\stackrel{(\text{A.18})}{=} \frac{1}{2} \cos^2(\gamma_1 - \gamma_2) \quad (\text{A.19})$$

→ zurück zu Gleichung (7.4)

A.4 Die Zustände $|\uparrow\rangle$ und $|\downarrow\rangle$ des Beryllium-Ions

Der **Drehimpuls** ist eine Erhaltungsgröße. Genau wie Energie und **Impuls** kann auch Drehimpuls nicht aus dem Nichts erscheinen und auch nicht im Nichts verschwinden. Ein Drehimpuls hat nicht nur einen Wert, sondern auch eine Richtung im Raum. Deshalb werden Drehimpulse in der Theorie durch Vektoren repräsentiert und mit Fettdruck gekennzeichnet. Wir werden im Folgenden für Drehimpulse den Buchstaben \mathbf{J} verwenden.

Früher glaubte man dass der Betrag eines Drehimpulses jeden beliebigen Wert haben könne. Erst bei der Untersuchung der Spektren von Atomen und Molekülen wurde entdeckt, dass der Betrag jedes Drehimpulses, ohne Ausnahme, einen der diskreten Werte

$$|\mathbf{J}| = \sqrt{j(j+1)} \hbar \quad \text{mit } j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \dots \quad (\text{A.20})$$

hat. In dieser Formel ist \hbar (sprich: ha quer) eine abkürzende Schreibweise für $h/(2\pi)$, wobei $h = 6,6 \cdot 10^{-34} \text{kg m}^2/\text{s}$ die Planck'sche Konstante und $\pi = 3,1415\dots$ das Verhältnis vom Umfang zum Durchmesser eines Kreises ist. j ist die Drehimpuls-Quantenzahl. Bei einem makroskopischen Kreisel ist die Quantenzahl j gigantisch groß, so dass man Änderungen von j nach $j \pm 1$ überhaupt nicht bemerkt, und der Drehimpuls eine kontinuierliche Größe zu sein scheint.

Auch der Spin von Elektronen und Atomkernen ist eine Form des Drehimpulses. Das Experiment von Stern und Gerlach (siehe Abschnitt 6.1) hat gezeigt, dass das magnetische Moment von Silberatomen in einem äußeren Magnetfeld nur zwei Richtungen haben kann. Weil dieses magnetische Moment mit dem Spin verbunden ist, bedeutet das Ergebnis von Stern und Gerlach zugleich dass der Drehimpuls eines Silberatoms relativ zum magnetischen Feld entweder die eine oder die andere von nur zwei möglichen Richtungen hat.

Im Lauf der folgenden Jahre zeigte die genauere quantitative Untersuchung, dass Elektronen die Spinquantenzahl $j = 1/2$ haben, und dass die Projektion dieses Spins auf die Richtung eines Magnetfeldes immer

$$\text{entweder } +\frac{\hbar}{2} \quad \text{oder} \quad -\frac{\hbar}{2} \quad (\text{A.21})$$

ist. Dies ist der Spezialfall einer allgemeinen Regel: Wenn die Projektion J_{\parallel} eines Drehimpulses mit Quantenzahl j auf eine bestimmte Richtung des Raums gemessen wird, dann hat sie stets einen der Werte

$$J_{\parallel} = j\hbar, (j-1)\hbar, (j-2)\hbar, \dots, -j\hbar. \quad (\text{A.22})$$

Beryllium existiert fast ausschließlich als stabiles Isotop ${}^9\text{Be}$. Sein Kern besteht aus 4 Protonen und 5 Neutronen. Das neutrale Atom hat vier Elektronen, das Be^+ Ion nur drei. Ein kleiner Ausschnitt aus dem Termschema des einfach positiv geladenen Be^+ Ions wird in Abb. A.4 gezeigt. Im Grundzustand (dem Zustand mit niedrigster Energie) hat die Elektronenhülle des Be^+ Ions die Drehimpuls-Quantenzahl $j_{\text{Elektronen}} = 1/2$. Die Drehimpuls-Quantenzahl des Atomkerns ist $j_{\text{Kern}} = 3/2$. Also kann die Drehimpuls-Quantenzahl des Atoms insgesamt $j = 3/2 + 1/2 = 2$ oder $j = 3/2 - 1/2 = 1$ sein. Die Projektion des Drehimpulses auf die Richtung eines Magnetfeldes kann nur einen der diskreten Werte $J_{\parallel} = (\text{A.22})$ annehmen. Im Fall des Be^+ Ions hat der Zustand mit $j = 2$ und $J_{\parallel} = -2\hbar$ die niedrigste Energie.

Im Zustandsvektor (7.12) sind die beiden Zustandsvektoren

$$|\uparrow\rangle = |j = 1, J_{\parallel} = -\hbar\rangle \quad (\text{A.23a})$$

$$|\downarrow\rangle = |j = 2, J_{\parallel} = -2\hbar\rangle \quad (\text{A.23b})$$

der einzelnen Be^+ Ionen miteinander verschränkt. Es handelt sich also um die jeweils untersten Zustände des Triplets ($j = 1$) und

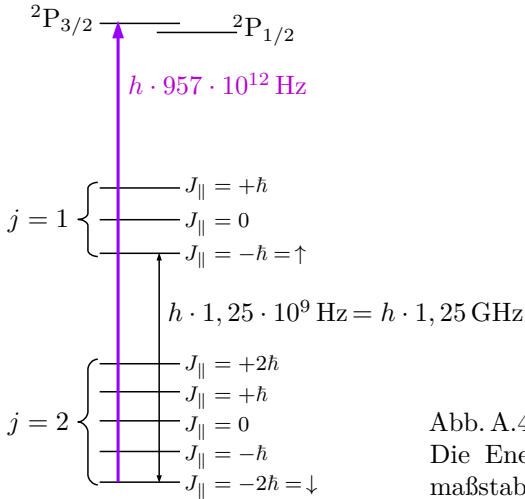


Abb. A.4: Term-Schema von ${}^9\text{Be}^+$. Die Energiedifferenzen sind nicht maßstabsgerecht gezeichnet!

des Quintetts ($j = 2$) im Termschema A.4. Man könnte gegen die Benennung \uparrow und \downarrow einwenden, dass doch die Projektionen $J_{||} = -\hbar$ und $J_{||} = -2\hbar$ beider Drehimpulse die gleiche Richtung haben. Aber genau wie im Gedankenexperiment von Bohm gilt auch im Experiment mit den zwei Be^+ Ionen

$$J_{||}(\uparrow) - J_{||}(\downarrow) = \hbar,$$

siehe das Termschema A.4. So gesehen ist es durchaus angemessen, die Schreibweise der Zustandsvektoren $|\uparrow\rangle$ und $|\downarrow\rangle$ auch jetzt wieder zu verwenden.

→ zurück nach Seite 172

A.5 Die Drehung der verschränkten Zustandsvektoren des Beryllium-Ions

Die Drehungen

$$|\uparrow\rangle_1 \xrightarrow{\text{Drehung um } \varphi_1} a_1 |\uparrow\rangle_1 + b_1 |\downarrow\rangle_1 \quad (\text{A.24a})$$

$$|\downarrow\rangle_1 \xrightarrow{\text{Drehung um } \varphi_1} c_1 |\uparrow\rangle_1 + d_1 |\downarrow\rangle_1 \quad (\text{A.24b})$$

$$|\uparrow\rangle_2 \xrightarrow{\text{Drehung um } \varphi_2} a_2 |\uparrow\rangle_2 + b_2 |\downarrow\rangle_2 \quad (\text{A.24c})$$

$$|\downarrow\rangle_2 \xrightarrow{\text{Drehung um } \varphi_2} c_2 |\uparrow\rangle_2 + d_2 |\downarrow\rangle_2, \quad (\text{A.24d})$$

der Zustände, in denen die Amplituden a_1, b_1, c_1, d_1 von einem variabel wählbaren Winkel φ_1 abhängen, und die Amplituden a_2, b_2, c_2, d_2 von einem variabel wählbaren Winkel φ_2 abhängen¹⁰⁰, führten Rowe et al. folgendermaßen aus:

Nachdem sie den verschränkten Zustand (7.14) erzeugt hatten, schossen sie einen weiteren kurzen Laserpuls auf die beiden Be^+ Ionen. Mit diesem Puls wurde den Ionen eine Phasenverschiebung aufgeprägt, wie in Abb. A.5 auf der nächsten Seite gezeigt. Die beiden türkisen Punkte symbolisieren die Positionen der beiden Ionen. In A.5(b) ist die Wellenlänge des Laserpulses etwas größer als in A.5(a), und in A.5(c) nochmals etwas größer. Dadurch erhält das Ion₂ in A.5(b) einen Phasenversatz von einer achte Wellenlänge, und in A.5(c) einen Phasenversatz von einer viertel Wellenlänge gegenüber A.5(a).

Eine zweite Möglichkeit zur Variation des Phasenversatzes besteht darin, die positive Spannung der vier äußeren kleinen Elektroden der Falle (siehe Abb. 7.6 auf Seite 172) noch weiter zu erhöhen. Dann werden die Ionen noch enger ins Zentrum der Falle gedrückt, und man erhält den in A.5(d) und A.5(e) skizzierten Phasenversatz.

¹⁰⁰ für Physiker: Rowe et al. definierten $a_j = c_j = \sqrt{1/2}$, $b_j = -i\sqrt{1/2}e^{-i\varphi_j}$, $d_j = -i\sqrt{1/2}e^{+i\varphi_j}$ für $j = 1$ und $j = 2$.






	φ_1	φ_2
(a) 	0	0
(b) 	0	$+45^\circ \hat{=} +\lambda/8$
(c) 	0	$+90^\circ \hat{=} +\lambda/4$
(d) 	$-45^\circ \hat{=} -\lambda/8$	$+135^\circ \hat{=} +3\lambda/8$
(e) 	$-90^\circ \hat{=} -\lambda/4$	$\pm 180^\circ \hat{=} \pm \lambda/2$

Fig. A.5 : Die Einstellung von φ_1 und φ_2

Da man zwei voneinander unabhängige Stellschrauben (Änderung der Wellenlänge des Lasers, Änderung der Elektrodenspannung) hat, kann man jede gewünschte Phasenverschiebung für Ion_1 und für Ion_2 einstellen.

Wenn man ein Objekt um 360° dreht, dann hat es genau die gleiche Stellung wie vor der Drehung. Eine Drehung um 360° hat also den gleichen Effekt wie eine Drehung um 0° . Und ein Phasenversatz von einer ganzen Wellenlänge hat genau den gleichen Effekt wie überhaupt kein Phasenversatz. In diesem Sinn entspricht (das Zeichen $\hat{=}$ bedeutet „entspricht“, und λ ist die Wellenlänge) eine Phasenverschiebung von $\pm\lambda$ einer Drehung um $\pm 360^\circ$, eine Phasenverschiebung von $\pm\lambda/2$ einer Drehung um $\pm 180^\circ$, eine Phasenverschiebung von $\lambda/8$ entspricht einer Drehung um 45° , und so weiter.

→ zurück nach Seite [176](#)

A.6 Das $\lambda/2$ -Plättchen

Ein $\lambda/2$ -Plättchen ist ein dünnes Plättchen eines [anisotropen](#) Kristalls (meist wird Glimmer verwendet), der von Licht mit unterschiedlicher Polarisation unterschiedlich schnell durchlaufen wird. Was dabei geschieht, kann man anhand von [Abb. A.2](#) auf Seite [334](#)

verstehen. Wenn das $\lambda/2$ -Plättchen so justiert ist dass seine „schnelle“ Achse in y -Richtung orientiert ist, dann läuft die gestrichelt gezeichnete Teilwelle mit L_0 -Polarisation schneller durch den Kristall als die gepunktet gezeichnete Teilwelle mit L_{90} -Polarisation. Wenn das Plättchen genau so dick geschnitten ist, dass die schnellere Teilwelle der langsameren Teilwelle am Ende des Kristalls gerade um eine halbe Wellenlänge voraus ist (daher hat das Plättchen seinen Namen), dann ist aus der L_{45} -Welle von Abb. A.2(a) die L_{135} -Welle von Abb. A.2(b) geworden. Das $\lambda/2$ -Plättchen hat die Polarisationsrichtung des Lichts also um 90° gedreht. (Man kann sich auch klarmachen, dass ein $\lambda/2$ -Plättchen das rechts-zirkular polarisierte Lichte von Abb. A.2(c) in das links-zirkular polarisierte Lichte von Abb. A.2(d) verwandelt, und umgekehrt.)

Wenn die schnelle Achse des $\lambda/2$ -Plättchens dagegen um 45° gegen die y -Achse gedreht ist, dann wird weder die Polarisation der L_{45} -Welle von Abb. A.2(a) noch die Polarisation der L_{135} -Welle von Abb. A.2(b) gedreht, weil diese Wellen dann entweder ausschließlich schnelleres Licht oder ausschließlich langsamer Licht enthalten. In diesem Fall gibt es keine schnellere Teilwelle, die eine langsamere Teilwelle überholen könnte.

Wenn das $\lambda/2$ -Plättchen je nach Einstellung die Polarisationsrichtung von Licht um 90° oder um 0° dreht, dann ist es plausibel – und man kann es sich mit etwas Überlegung auch detailliert klarmachen – dass man die Polarisationsrichtung von linear polarisiertem Licht durch geeignete Justierung des $\lambda/2$ -Plättchens um jeden beliebigen Winkel drehen kann.

→ zurück nach Seite 186

A.7 The marking of Rubidium atoms

In its ground state, the electron shell of the rubidium atom has an angular momentum quantum number $j_e = 1/2$, and the nucleus of

the rubidium-isotope ^{85}Rb used in this experiment has the angular momentum quantum number $j_{\text{N}} = 5/2$. These two quantum numbers can be combined to the total angular-momentum quantum number $j = 5/2 + 1/2 = 3$ or $j = 5/2 - 1/2 = 2$. The state with $j = 2$ is the ground state (i. e. the state with the lowest energy); the energy of the state with $j = 3$ is $h \cdot 3.04 \text{ GHz}$ higher.¹⁰¹

We define this notation for the state vector of the Rubidium atoms:

$$|2\rangle = \text{ground state with } j = 2 \quad (\text{A.25a})$$

$$|3\rangle = \text{excited state with } j = 3 \text{ and} \\ \text{energy } h \cdot 3.04 \text{ GHz above ground state} \quad (\text{A.25b})$$

If a rubidium atom, which is initially in state $|2\rangle$, is irradiated with microwaves of frequency 3.04 GHz, then it oscillates back and forth between the states $|2\rangle$ and $|3\rangle$,¹⁰² and it oscillates the faster, the higher the intensity of the microwave radiation. A microwave pulse that excites the atom with probability 1/2 from state $|2\rangle$ to state $|3\rangle$ is called a $\pi/2$ -pulse (pronounced: pi-half pulse):

$$|2\rangle \xrightarrow{\pi/2\text{-pulse}} \frac{1}{\sqrt{2}} \left(|2\rangle + i|3\rangle \right) \quad (\text{A.26a})$$

I will not explain the phase factor $i = \sqrt{-1}$, as that would go beyond the scope of this book.¹⁰³ The factor $\sqrt{1/2}$ was inserted so that the projection of this state onto itself is equal to 1. Let's

¹⁰¹ Each of these energy levels splits further into sublevels in an external magnetic field, as described in fig. A.4 on page 347 for the experiment with beryllium ions. This splitting, however, is irrelevant for the marking of the rubidium atoms.

¹⁰² for physicists: An elementary introduction to the theory of Rabi-oscillations can be found in [48].

¹⁰³ For physicists: These are the same phase factors as in (A.8). See also footnote 102.

verify this (as always, the sign of i must be reversed in the left-hand side of the projection amplitude):

$$\begin{aligned} & \frac{1}{\sqrt{2}} \left(\langle 2| - i\langle 3| \right) \frac{1}{\sqrt{2}} \left(|2\rangle + i|3\rangle \right) = \\ & = \frac{1}{2} \left(\underbrace{\langle 2||2\rangle}_1 + i \underbrace{\langle 2||3\rangle}_0 - i \underbrace{\langle 3||2\rangle}_0 \underbrace{-i^2}_{+1} \underbrace{\langle 3||3\rangle}_1 \right) = 1 \end{aligned}$$

If the atom were irradiated with three additional $\pi/2$ -pulses, its state vector would change as follows:

$$\frac{1}{\sqrt{2}} \left(|2\rangle + i|3\rangle \right) \xrightarrow{\pi/2\text{-pulse}} i|3\rangle \quad (\text{A.26b})$$

$$i|3\rangle \xrightarrow{\pi/2\text{-pulse}} \frac{1}{\sqrt{2}} \left(-|2\rangle + i|3\rangle \right) \quad (\text{A.26c})$$

$$\frac{1}{\sqrt{2}} \left(-|2\rangle + i|3\rangle \right) \xrightarrow{\pi/2\text{-pulse}} -|2\rangle \quad (\text{A.26d})$$

Here, too, I will not explain the phase factors ± 1 and $\pm i$.¹⁰³ Before the first $\pi/2$ -pulse, the atom was in state $|2\rangle$ with probability 1; see the left-hand side of (A.26a). After the first $\pi/2$ -pulse, it is with probability 1/2 in state $|2\rangle$, and with probability 1/2 in state $|3\rangle$, see the right-hand side of (A.26a).

After the second $\pi/2$ -pulse, the atom is with probability 1 in state $|3\rangle$; see the right-hand side of (A.26b). After the third $\pi/2$ -pulse, it is with probability 1/2 in state $|2\rangle$, and with probability 1/2 in state $|3\rangle$, see the right-hand side of (A.26c). After the fourth $\pi/2$ -pulse, it is with probability 1 finally back in the state $|2\rangle$, see the right-hand side of (A.26d).

The following fact, which again I will not explain, is important for marking the atoms:¹⁰⁴ If an atom in state $|2\rangle$ is reflected, then

¹⁰⁴ Physicists can find the explanation in [58]. Incidentally, the authors later changed their minds and wrote in more recent publications that the trans-

the sign of the state vector changes. If the atom is in state $|3\rangle$, the state vector remains unchanged upon reflection:

$$|2\rangle \xrightarrow{\text{reflection due to laser-wave}} -|2\rangle \quad (\text{A.27a})$$

$$|3\rangle \xrightarrow{\text{reflection due to laser-wave}} +|3\rangle \quad (\text{A.27b})$$

By skillfully combining (A.26) and (A.27), Dürr et al. were able to mark the rubidium atoms as follows:

Initially, as they fall through slit S2 from above, the atoms are in the $|2\rangle$ state. Then — even before the laser field, which acts as a beam splitter, was switched on — they were irradiated with a $\pi/2$ -pulse of the 3.04 GHz-microwave field, and thereby excited according to (A.26a) to state

$$\frac{1}{\sqrt{2}} \left(|2\rangle + i|3\rangle \right). \quad (\text{A.28a})$$

The atomic beam was then split by the first laser field, which was turned on for $45 \mu\text{s}$. If an atom is reflected and takes path B, then according to (A.27), the sign of $|2\rangle$ changes, but the sign of $|3\rangle$ does not. If the atom is transmitted and takes path A, the signs of $|2\rangle$ and $|3\rangle$ remain unchanged. Thus, the state vector of the atoms after the first beam splitting was

$$\frac{1}{\sqrt{2}} \left[\underbrace{\frac{1}{\sqrt{2}} \left(|2\rangle + i|3\rangle \right)}_{\text{path A}} + \underbrace{\frac{1}{\sqrt{2}} \left(-|2\rangle + i|3\rangle \right)}_{\text{path B}} \right]. \quad (\text{A.28b})$$

The factor $\sqrt{1/2}$ in front of the square brackets was added so that the projection of the state vector onto itself is equal to 1.

mitted state $|2\rangle$ changes its sign, but the reflected state $|2\rangle$ does not. This seems quite dubious to me, and in this experiment it doesn't matter in which beam the sign change occurs.

$|2\rangle$ or $|3\rangle$ is a highly abbreviated notation for the state vectors of the atoms. In fact, we know much more than just that the atoms have angular momentum quantum numbers $j = 2$ or $j = 3$. We know that these are rubidium atoms, we know the path they are traveling along, we know their velocity, and much more. It will soon prove useful to write the state vectors as products

$$\begin{aligned} |2\rangle &\longleftrightarrow |2\rangle|\text{all other informations}\rangle \\ |3\rangle &\longleftrightarrow |3\rangle|\text{all other informations}\rangle . \end{aligned}$$

The first factor contains nothing more than the information about the angular momentum quantum number; the rest of the complete state vector is packed into the second factor. Specifically, we include for the moment being into the second factor the path along which the atom is traveling:

$$\begin{aligned} \text{(A.28b)} &= \frac{1}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} \left(|2\rangle|A\rangle + i|3\rangle|A\rangle \right) + \frac{1}{\sqrt{2}} \left(-|2\rangle|B\rangle + i|3\rangle|B\rangle \right) \right] \\ &\hspace{15em} \text{(A.28c)} \end{aligned}$$

The atoms were then irradiated with a second $\pi/2$ -pulse of the microwave field (3.04 GHz). From (A.26b) and (A.26d), we can see that subsequently the state vector of the atoms was

$$\frac{1}{\sqrt{2}} \left[i|3\rangle|A\rangle - |2\rangle|B\rangle \right] . \quad \text{(A.28d)}$$

Finally, the beam-splitting laser field was switched on for a second time for $45 \mu\text{s}$, so that the atoms with probability $1/2$ were reflected. Thereby their state vector became

$$\begin{aligned} &\frac{1}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} \left(i|3\rangle|H_A\rangle + i|3\rangle|G_A\rangle \right) + \frac{1}{\sqrt{2}} \left(-|2\rangle|G_B\rangle + |2\rangle|H_B\rangle \right) \right] \\ &= \frac{1}{2} \left(i|3\rangle|G_A\rangle - |2\rangle|G_B\rangle + i|3\rangle|H_A\rangle + |2\rangle|H_B\rangle \right) . \quad \text{(A.28e)} \end{aligned}$$

Here it was noted that according to (A.27) the sign of $|2\rangle$ changes upon reflection, whereas that of $|3\rangle$ does not.

If the atoms are *not* marked by microwave pulses, instead of (A.28) the following sequence of state vectors is obtained:

Before the first reflection, the state vector of the atoms is

$$|2\rangle . \quad (\text{A.29a})$$

Then the atomic beam is split by the first laser field, which is turned on for $45\ \mu\text{s}$. If an atom is reflected and takes path B, then according to (A.27) the sign of $|2\rangle$ changes; if the atom is transmitted and takes path A, however, then the sign does not change. Consequently the state vector of the atoms after the first beam split is

$$\frac{1}{\sqrt{2}} \left(|2\rangle|A\rangle - |2\rangle|B\rangle \right) . \quad (\text{A.29b})$$

After the beam-splitting laser field has been turned-on a second time for $45\ \mu\text{s}$, the state vector of the atoms eventually is







$$\begin{aligned} & \frac{1}{\sqrt{2}} \left[|2\rangle \frac{1}{\sqrt{2}} \left(|H_A\rangle - |G_A\rangle \right) - |2\rangle \frac{1}{\sqrt{2}} \left(|G_B\rangle - |H_B\rangle \right) \right] = \\ & = \frac{1}{2} \left(-|2\rangle|G_A\rangle - |2\rangle|G_B\rangle + |2\rangle|H_A\rangle + |2\rangle|H_B\rangle \right) . \quad (\text{A.29c}) \end{aligned}$$










→ back to page [211](#)









References









 pay-wall  open access








Some of the free links are “unstable”, i. e. the documents sometimes disappear from the servers after a few years, or the entire website is taken down. In such cases, search engines can often find the document on another (free of charge) server if you search for the title and the author’s name. If this problem occurs, please let me know. [mailto: gerold.gruendler@astrophys-neunhof.de](mailto:gerold.gruendler@astrophys-neunhof.de)











- [1] Claudius Ptolemaeus: *Syntaxis Mathematica* Vol. I & II (B. G. Teubner, Leipzig, 1898) Greek text, edited by J. L. Heiberg. Volume I:  <http://www.wilbourhall.org/millionbookspdfs/pt1claudiptolemaei01ptoluoft.pdf>
Volume I & II:  <http://www.wilbourhall.org/pdfs/HeibergAlmagestComplete.pdf> English translation by G. J. Toomer (Duckworth & Co. Ltd., London, 1984):  https://classicalliberalarts.com/resources/PTOLEMY_ALMAGEST_ENGLISH.pdf
- [2] Isaac Newton: *Philosophiae Naturalis Principia Mathematica* (Guil. & Joh. Innys, London, 3rd ed. 1726)
 <https://dx.doi.org/10.3931/e-rara-1235>, English translation:  <https://archive.org/details/newtonspmathema00newtrich> or  [https://en.wikisource.org/wiki/The_Mathematical_Principles_of_Natural_Philosophy_\(1846\)](https://en.wikisource.org/wiki/The_Mathematical_Principles_of_Natural_Philosophy_(1846))










- [3] Pierre Simon Laplace: *Essai philosophique sur les probabilités* (1814)  <https://archive.org/details/essaiphilosophiq00lapluoft/mode/2up>
English translation: *A philosophical essay on probabilities*  <https://archive.org/details/cu31924001150733>
- [4] *Evanescent Electromagnetic Fields*
APIN Circular se91013 (2014)
 <https://astrophys-neunhof.de/mtlg/se91013.pdf>
- [5] F. Hénault: *Quantum physics and the beam splitter mystery*,
Proc. SPIE 9570, VI, 95700Q (2015),
 <https://dx.doi.org/10.1117/12.2186291>
 <https://arxiv.org/abs/1509.00393>
- [6] Thomas Young: *Experiments and Calculations Relative to Physical Optics*,
Phil. Trans. Roy. Soc. London **94**, 1–16 (1804)
 <https://www.jstor.org/stable/pdf/107135.pdf>
- [7] P. Lenard: *Ueber die lichtelektrische Wirkung*,
Ann. Phys. (Leipzig) **8**, 149–198 +Taf.I (1902),
 <https://dx.doi.org/10.1002/andp.19023130510>
 https://grundpraktikum.physik.uni-saarland.de/gpalt/Anleitungen/Ergaenzungen/J1_Papers/Photoeffekt%20-%20Lenard_1.pdf
- [8] A. Einstein: *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*,
Ann. Phys. (Leipzig) **17**, 132–148 (1905)
 <https://dx.doi.org/10.1002/andp.19053220607>
English translation:  https://www.informationphilosopher.com/solutions/scientists/einstein/AJP_1905_photon.pdf









- [9] A. Einstein: *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*,
Ann. Phys. (Leipzig) **17**, 549 – 560 (1905),
 <https://dx.doi.org/10.1002/andp.19053220806>
- [10] A. Einstein: *Zur Elektrodynamik bewegter Körper*,
Ann. Phys. (Leipzig) **17**, 891 – 920 (1905),
 <https://dx.doi.org/10.1002/andp.19053221004>
- [11] Max Planck: *Die Entstehung und bisherige Entwicklung der Quantentheorie*, Nobelvortrag 2. Juni 1920
(J. A. Barth, Leipzig, 1920)
English translation:  https://nobelprize.org/nobel_prizes/physics/laureates/1918/planck-lecture.html
- [12] R. A. Millikan: *A direct photoelectric determination of Planck's h* , Phys. Rev. **7**, 355 – 388 (1916),
 <https://dx.doi.org/10.1103/PhysRev.7.355>
- [13] A. H. Compton: *A Quantum Theory of the Scattering of X-rays by Light Elements*, Phys. Rev. **21**, 483 – 502 (1923),
 <https://dx.doi.org/10.1103/PhysRev.21.483>
- [14] G. I. Taylor: *Interference fringes with feeble light*,
Proc. Cambridge Phil. Soc. **15**, 114 – 115 (1909),  <https://www.biodiversitylibrary.org/item/97262#page/142/mode/1up>
- [15] J. J. Thorn, M. S. Neel, V. W. Donato, G. S. Bergreen,
R. E. Davies, M. Beck: *Observing the quantum behavior of light in an undergraduate laboratory*,
Am. J. Phys. **72**, 1210 – 1219 (2004),
 <https://dx.doi.org/10.1119/1.1737397>  https://www.reed.edu/physics/332/pdf/Single_Photons.pdf

- [16] E. J. Galvez, C. H. Holbrow, M. J. Pysher, J. W. Martin, N. Courtemanche, L. Heilig, J. Spencer: *Interference with correlated photons: Five quantum mechanics experiments for undergraduates*, Am. J. Phys. **73**, 127 – 140 (2005),
 <https://dx.doi.org/10.1119/1.1796811>
 <https://egalvez.colgate.domains/pql/wp-content/uploads/2019/08/ajpph.pdf>
- [17] Christian Morgenstern: *Alle Galgenlieder* (Insel Verlag, Frankfurt am Main, 1972)
 <https://projekt-gutenberg.org/authors/christian-morgenstern/books/alle-galgenlieder>
- [18] A. Zeilinger, R. Gähler, C. G. Shull, W. Treimer, W. Mampe: *Single- and double-slit diffraction of neutrons*, Rev. Mod. Phys. **60**, 1067 – 1073 (1988),
 <https://dx.doi.org/10.1103/RevModPhys.60.1067>
 <https://members.ift.uam-csic.es/bellido/cuantica/articulos/RevModPhys.60.1067.pdf>
- [19] S. Frabboni, G. C. Gazzadi, G. Pozzi: *Nanofabrication and the realization of Feynman's two-slit experiment*, Appl. Phys. Lett. **93**, 073108 (3pp) (2008),
 <https://dx.doi.org/10.1063/1.2962987>
 [https://wiki.epfl.ch/mep/documents/MEP\[08-09\]_DOWNLOAD/applphyslett_93_073108_feynman_exp.pdf](https://wiki.epfl.ch/mep/documents/MEP[08-09]_DOWNLOAD/applphyslett_93_073108_feynman_exp.pdf)
- [20] R. Bach, D. Pope, S. H. Liou, H. Batelaan: *Controlled double-slit electron diffraction*, New J. Phys. **15**, 033018 (7pp) (2013),
 <https://dx.doi.org/10.1088/1367-2630/15/3/033018>









- [21] R. Bach, D. Pope, S. H. Liou, H. Batelaan :
Supplement to: Controlled double-slit electron diffraction,
New J. Phys. **15**, (2013),  <https://iopscience.iop.org/article/10.1088/1367-2630/15/3/033018/data>
- [22] O. Carnal, J. Mlynek : *Young's double-slit experiment with atoms: A simple atom interferometer*,
Phys. Rev. Lett. **66**, 2689 (1991),
 <https://dx.doi.org/10.1103/PhysRevLett.66.2689>
- [23] F. Shimizu, K. Shimizu, H. Takuma : *Double-slit interference with ultracold metastable neon atoms*,
Phys. Rev. A **46**, R17 – R20 (1992),
 <https://dx.doi.org/10.1103/PhysRevA.46.R17> or:  http://ressources.agreg.phys.ens.fr/static/Videos/Shimuzu/92PRA%20Double-slit%20interference%20with%20ultracold%20metastable%20neon%20atoms_Shimizu.pdf
- [24] Galileo Galilei : *Discorsi e dimostrazioni matematiche intorno a due nuove scienze* (Elsevier, Leiden, 1638),
 <https://liberliber.it/autori/autori-g/galileo-galilei/discorsi-e-dimostrazioni-matematiche-intorno-a-due-nuove-scienze/>
english translation: *Dialogues Concerning Two New Sciences* (Dover Publ., New York, 1954),  http://galileoandeinstein.physics.virginia.edu/tns_draft/index.html
deutsche Übersetzung: *Unterredungen und mathematische Demonstrationen* (W. Engelmann, Leipzig, 1890),
 https://openlibrary.org/works/OL5458481W/Unterredungen_und_mathematische_Demonstrationen


- [25] M. S. Chapman, C. R. Ekstrom, T.D. Hammond, R.A. Rubenstein, J. Schmiedmayer, S. Wehinger, D. E. Pritchard : *Optics and Interferometry with Na₂ Molecules*, Phys. Rev. Lett. **74**, 4783–4786 (1995),
 <https://dx.doi.org/10.1103/PhysRevLett.74.478> or:
 http://chapmanlabs.gatech.edu/papers/molecular_ifm_prl95.pdf
- [26] K. Hornberger, S. Gerlich, P. Haslinger, S. Nimmrichter, M. Arndt : *Colloquium: Quantum interference of clusters and molecules*, Rev. Mod. Phys. **84**, 157–173 (2012),
 <https://dx.doi.org/10.1103/RevModPhys.84.157>
 <https://doi.org/10.48550/arXiv.1109.5937>
- [27] S. Eibenberger, S. Gerlich, M. Arndt, M. Mayor, J. Tüxen : *Matter-wave interference of particles selected from a molecular library with masses exceeding 10 000 amu*, Phys. Chem. Chem. Phys. **15**, 14696–14700 (2013),
 <https://doi.org/10.1039/C3CP51500A>
- [28] Galileo Galilei : *Il Saggiatore* (1623)  <https://liberliber.it/autori/autori-g/galileo-galilei/il-saggiatore/>
english translation: *The Assayer*  <https://web.stanford.edu/~jsabol/certainty/readings/Galileo-Assayer.pdf>
- [29] M. Born : *Zur Quantenmechanik der Stoßvorgänge*, Zeits. Phys. **37**, 863–867 (1926),  <https://dx.doi.org/10.1007/BF01397477> oder:  http://www.psiquadra.t.de/downloads/born26_stossvorgaenge.pdf,
english translation: [76, pp. 52–55]
- [30] N. Bohr : *On the Constitution of Atoms and Molecules*, Philos. Mag. **26**, 1–24 (1913)  <https://www.bibnum.education.fr/sites/default/files/bohr-texte.pdf>










- [31] Werner Heisenberg: *Der Teil und das Ganze*
(Piper Verlag, München, 1969)
english translation: *Physics and Beyond*
(Harper & Row, New York, USA, 1971)
- [32] E. Schrödinger: *Discussion of Probability Relations between Separated Systems*,
Math. Proc. Cambridge Phil. Soc. **31**, 555–563 (1935),
 <https://dx.doi.org/10.1017/S0305004100013554> or:
 <http://www.informationphilosopher.com/solutions/scientists/schrodinger/Schrodinger-1935.pdf>
- [33] W. Gerlach, O. Stern: *Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld*,
Z. Phys. **9**, 349–352 (1922),
 <https://dx.doi.org/10.1007/BF01326983>
english translation:
 <https://doi.org/10.48550/arXiv.2301.11343>
- [34] P. A. M. Dirac: *The Quantum Theory of the Electron*,
Proc. Roy. Soc. (London) A **117**, 610–624 (1928)
 <https://doi.org/10.1098/rspa.1928.0023>
- [35] P. A. M. Dirac: *The Quantum Theory of the Electron. Part II*,
Proc. Roy. Soc. (London) A **118**, 351–624 (1928)
 <https://dx.doi.org/10.1098/rspa.1928.0056>
- [36] W. Heisenberg: *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*,
Z. Phys. **43**, 172–198 (1927),
 <https://dx.doi.org/10.1007/BF01397280>  <http://fisicafundamental.net/relicario/doc/Heisenberg1927.pdf>
english translation:
 <https://ntrs.nasa.gov/citations/19840008978>

- [37] A. Einstein, B. Podolsky, N. Rosen: *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*, Phys. Rev. **47**, 777–780 (1935)
 <https://dx.doi.org/10.1103/PhysRev.47.777>
- [38] David Bohm: *Quantum Theory* (Prentice-Hall, Englewood Cliffs, NJ, USA, 1951)
- [39] N. Bohr: *Can Quantum-Mechanical Description of Physical Reality be Considered Complete?*, Phys. Rev. **48**, 696–702 (1935)
 <https://dx.doi.org/10.1103/PhysRev.48.696>
- [40] J. S. Bell: *On the Einstein Podolski Rosen Paradox*, Physics **1**, 195–200 (1964)  https://cds.cern.ch/record/111654/files/vol1p195-200_001.pdf
- [41] A. Peres: *Unperformed experiments have no results*, Am. J. Phys. **46**, 745–747 (1978),
 <https://dx.doi.org/10.1119/1.11393>
 https://www.if.ufrj.br/~pef/aulas_seminarios/notas_de_aula/carlos_2012_1/seminariosTemasMQ/teoremaBell/teoremaBell_Peris_AJP1978.pdf
- [42] J. S. Bell: *The Theory of local Beables*, CERN Report TH-2053, 14 pp (1975)
 <https://cds.cern.ch/record/980036/files/197508125.pdf>
- [43] J. G. Cramer: *Generalized absorber theory and the Einstein-Podolsky-Rosen paradox*, Phys. Rev. D **22**, 362–376 (1980)
 <https://doi.org/10.1103/PhysRevD.22.362>  https://faculty.washington.edu/jcramer/TI/PRD_22_362_1980.pdf

- [44] R. E. Kastner : *The Transactional Interpretation and its Evolution into the 21st Century: An Overview*,
Philos. Comp. **11**, 923 – 932 (2016),
 <https://dx.doi.org/10.1111/phc3.12360>
 <https://doi.org/10.48550/arXiv.1608.00660>
- [45] A. Aspect, P. Grangier, G. Roger : *Experimental Tests of Realistic Local Theories via Bell's Theorem*,
Phys. Rev. Lett. **47**, 460 – 463 (1981),
 <https://dx.doi.org/10.1103/PhysRevLett.47.460>
- [46] A. Aspect, P. Grangier, G. Roger : *Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities*,
Phys. Rev. Lett. **49**, 91 – 94 (1982),
 <https://dx.doi.org/10.1103/PhysRevLett.49.91>
- [47] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett,
W. M. Itano, C. Monroe, D. J. Wineland : *Experimental violation of a Bell's inequality with efficient detection*,
Nature **409**, 791 – 794 (2001),
 <https://dx.doi.org/10.1038/35057215>
 <https://backend.production.deepblue-documents.lib.umich.edu/server/api/core/bitstreams/6fc52f2d-df2f-4b27-a18f-2c40f1a0a37c/content>
- [48] *Interaction of 2-level systems and electromagnetic radiation*
APIN Circular se14211 (2016)
 <https://www.astrophys-neunhof.de/mtlg/se14211.pdf>

- [49] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, Y. Shih: *New High-Intensity Source of Polarization-Entangled Photon Pairs*, Phys. Rev. Lett. **75**, 4337–4341 (1995),
 <https://dx.doi.org/10.1103/PhysRevLett.113.120405>
 [https://research.physics.illinois.edu/QI/Photonics/papers/My Collection.Data/PDF/New high-intensity source of polarization-entangled photon pairs.pdf](https://research.physics.illinois.edu/QI/Photonics/papers/My%20Collection/Data/PDF/New%20high-intensity%20source%20of%20polarization-entangled%20photon%20pairs.pdf)
- [50] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, A. Zeilinger: *Violation of Bell's Inequality under Strict Einstein Locality Conditions*, Phys. Rev. Lett. **81**, 5039–5043 (1998),
 <https://doi.org/10.1103/PhysRevLett.81.5039>
 <https://doi.org/10.48550/arXiv.quant-ph/9810080>
- [51] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenber, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, R. Hanson: *Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres*, Nature **526**, 682–686 (2015)
 <https://dx.doi.org/10.1038/nature15759>
 <https://doi.org/10.48550/arXiv.1508.05949>
- [52] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J. A. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, A. Zeilinger: *Significant-loophole-free test of Bell's theorem with entangled photons*, Phys. Rev. Lett. **115**, 250401 (2015)
 <https://dx.doi.org/10.1103/PhysRevLett.115.250401>
 <https://doi.org/10.48550/arXiv.1511.03190>

- [53] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, S. W. Nam: *A strong loophole-free test of local realism*, Phys. Rev. Lett. **115**, 250401 (2015)
 <https://dx.doi.org/10.1103/PhysRevLett.115.250402>
 <https://doi.org/10.48550/arXiv.1511.03189>
- [54] W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, H. Weinfurter: *Event-ready Bell-test using entangled atoms simultaneously closing detection and locality loopholes*, Phys. Rev. Lett. **119**, 010402 (2017)
 <https://dx.doi.org/10.1103/PhysRevLett.119.010402>
- [55] A. Pais: *Einstein and the quantum theory*, Rev. Mod. Phys. **51**, 863 – 914 (1979),
 <https://dx.doi.org/10.1103/RevModPhys.51.863>
 <http://rpddata.caltech.edu/courses/aph105c/2006/articles/Pais.pdf>  https://eclass.aegean.gr/modules/document/file.php/511165/projects/einstein_quantum.pais.pdf
- [56] Max Planck: *Wissenschaftliche Selbstbiographie* (Johann Ambrosius Barth, Leipzig, 1948)
Auszug daraus in: Phys. Blätter **14**, 145 – 152 (1958),
 <https://dx.doi.org/10.1002/phbl.19580140401>
- [57] *Robertson's derivation of the indeterminacy relation*, APIN Circular se95311 (2010)
 <https://astrophys-neunhof.de/mtlg/se95311.pdf>

- [58] S. Dürr, T. Nonn, G. Rempe: *Origin of quantum-mechanical complementarity probed by a ‘which-way’ experiment in an atom interferometer*, Nature **395**, 33 – 37 (1998),
 <https://dx.doi.org/10.1038/25653>
- [59] S. Kunze, S. Dürr, G. Rempe: *Bragg scattering of slow atoms from a standing light wave*,
Europhys. Lett. **34**, 343 – 348 (1996),
 <https://doi.org/10.1209/epl/i1996-00462-x>
- [60] Y.-H. Kim, R. Yu, S. P. Kulik, Y. Shih, M. O. Scully: *Delayed “Choice” Quantum Eraser*, Phys. Rev. Lett. **84**, 1 – 5 (2000),
 <https://doi.org/10.1103/PhysRevLett.84.1>
 <https://doi.org/10.48550/arXiv.quant-ph/9903047>
- [61] M. O. Scully, K. Drühl: *Quantum eraser: A proposed photon correlation experiment concerning observation and “delayed choice” in quantum mechanics*,
Phys. Rev. A **25**, 2208 – 2213 (1982),
 <https://doi.org/10.1103/PhysRevA.25.2208>
- [62] A. C. Elitzur, L. Vaidman: *Quantum mechanical interaction-free measurements*,
Found. Phys. **23**, 987 – 997 (1993),
 <https://dx.doi.org/10.1007/BF00736012>
 <https://doi.org/10.48550/arXiv.hep-th/9305002>
- [63] L. Vaidman: *The Meaning of the Interaction-Free Measurements*, Found. Phys. **33**, 491 – 510 (2003),
 <https://dx.doi.org/10.1023/A:1023767716236>
 <https://doi.org/10.48550/arXiv.quant-ph/0103081>

- [64] P. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, M. A. Kasevich: *Interaction-Free Measurement*, Phys. Rev. Lett. **74**, 4763–4766 (1995),
 <https://doi.org/10.1103/PhysRevLett.74.4763>  <https://research.physics.illinois.edu/QI/Photonics/papers/My%20Collection.Data/PDF/Interaction-Free%20Measurement.pdf>
- [65] Werner Heisenberg: *The Copenhagen Interpretation of Quantum Theory*, in *Physics and Philosophy* Gifford Lectures, Univ. St. Andrews, Scotland, 1955/56 (Harper & Brothers, New York, USA, 1958),  <https://www.astrophys-neunhof.de/serv/Heisenberg1955.pdf>
- [66] J. N. Tinsley, M. I. Molodtsov, R. Prevedel, D. Wartmann, J. Espigulé-Pons, M. Lauwers, A. Vaziri: *Direct detection of a single photon by humans*, Nature Com. **7**, 12172 (2016)
 <https://dx.doi.org/10.1038/ncomms12172>
- [67] *Möglichkeit und Form*, APIN Mitteilung sd00211 (2011)
 <https://www.astrophys-neunhof.de/mtlg/sd00211.pdf>
- [68] Johann v. Neumann: *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932)
 <https://gdz.sub.uni-goettingen.de/id/PPN379400774>
 english translation: *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955)
 <https://ia600101.us.archive.org/11/items/mathematical-foundations-of-quantum-mechanics-2018/Mathematical%20Foundations%20of%20Quantum%20Mechanics%202018.pdf>
- [69] J. S. Bell: *Against ‘Measurement’*, Physics World **3**, 33–40 (Aug. 1990)
 <http://stacks.iop.org/2058-7058/3/i=8/a=26>  <http://www.tau.ac.il/~quantum/Vaidman/IQM/BellAM.pdf>

- [70] *Decoherence*, APIN Circular se91319 (2013)
 <https://astrophys-neunhof.de/mtlg/se91319.pdf>
- [71] L. Hackermüller, K. Hornberger, B. Brezger, A. Zeilinger, M. Arndt: *Decoherence of matter waves by thermal emission of radiation*, *Nature* **427**, 711 – 714 (2004),
 <https://dx.doi.org/10.1038/nature02276>
arXiv quant-ph/0402146 (2004)
 <https://arxiv.org/abs/quant-ph/0402146>
- [72] L. Boltzmann: *Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht*. *Wien. Ber.* **76**, 373 – 435 (1877)
auch in: L. Boltzmann: *Wissenschaftliche Abhandlungen* Band II (J. A. Barth, Leipzig, 1909)
 <https://phaidra.univie.ac.at/view/o:63651>
- [73] *Die Entdeckung der Bose-Einstein-Statistik*, APIN Mitteilung sd99331 (2011)
 <https://www.astrophys-neunhof.de/mtlg/sd99331.pdf>
- [74] A. Einstein: *Strahlungs-Emission und -Absorption nach der Quantentheorie*, *Verh. Deutsch. Phys. Ges.* **18**, 318 – 323 (1916)
 ausführlich zitiert in [75, Abs.3]
English translation:  https://www.informationphilosophie.com/solutions/scientists/einstein/1916_A-B.html
- [75] *Quantensprünge*, (APIN, Nürnberg, 2010)
 <https://www.astrophys-neunhof.de/mtlg/sd67113.pdf>
- [76] J. A. Wheeler, W. H. Zurek (Editors):
Quantum theory and measurement
(Princeton University Press, Princeton, New Jersey, 1983)