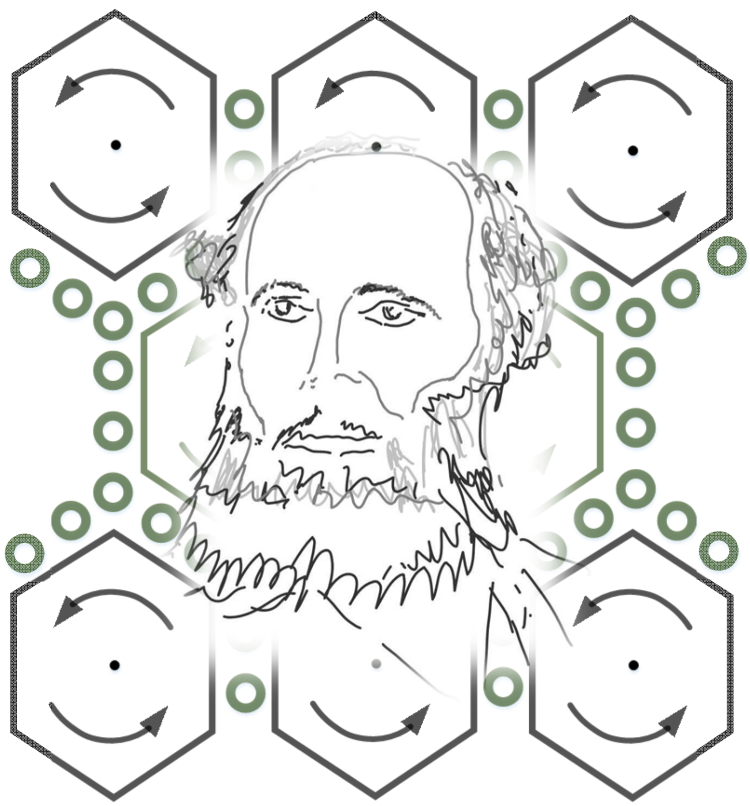
Quantum Information Theory for Engineers: **Free Climbing through Physics and Probability**

Supplement written by Christian Jansson

June 12, 2019



Contents

1 Introduction

The true logic of the world is in the calculus of probabilities.
James Clerk Maxwell

It is believed that quantum mechanics is the fundamental physical theory. Most physicists, not all, believe that it is a probabilistic theory describing microscopic systems. Thus, in the words of Maxwell, quantum theory is (perhaps) the true logic of the world.

The concept of probability is related to phenomena with several uncertain outcomes, the latter forming mutually exclusive alternatives. According to the Cambridge dictionary, a probability is a number that represents how likely it is that a particular outcome will happen. In other words, a probability describes a quantitative measure of the uncertainty of an outcome.

In many cases, by measuring the relative frequencies of the occurrence of certain outcomes, there should be no difficulty in empirically testing probabilities that have been predicted theoretically. But the meaning of the mathematical concepts of probability is by no means predetermined. Today, several interpretations of probability are discussed extensively. Already von Weizsäcker[1] wrote:

> *Probability is one of the outstanding examples of the epistemological paradox that we can successfully use our basic concepts without actually understanding them.* von Weizsäcker

Numerical probabilities don't come out of nothing. They don't arise out of a measure theory of probability, that is, they don't occur from mathematical axioms of a certain probability measure on a set of outcomes, like in Kolmogorov's classical probability theory. Historically, the first principle to get numerical probabilities was achieved by determining a set of mutually exclusive elementary events, the outcomes, such that there was no reason to discriminate. Then, these elementary events were assigned probabilities with the same positive values, all summing up to one. This approach is known as the *principle of indifference.*

The right way which to assign probabilities to the elementary events is a controversial philosophical discussion. Thus, for a better understanding, we shall investigate the following questions concerning probabilities:

Formal aspect: Is there a widely accepted definition of probability?

Variety: What sorts of things are probabilities?

Rules: Are there universal mathematical rules or axioms that can be used in **all** applications, from coin tossing to quantum electrodynamics?

Time: Are probabilities time dependent, and if so, in what form?

Quantum Probability: What is the relationship between classical probability and quantum probability?

[1] von Weizsäcker [2006, pp. 59]

It's worth reading Einstein's article[2] about physics and reality:

> *The aim of science is, on the one hand, a comprehension, as complete as possible, of the connection between the sense experiences in their totality, and, on the other hand, the accomplishment of this aim* **by the use of a minimum of primary concepts and relations.** Einstein 1936

Bohr said:

> *There is no quantum world. There is only an abstract quantum physical description. It is wrong to think that the task of physics is to find out how nature is. Physics concerns what we can say about nature.* Bohr[3]

Nowadays, the fundamental concepts of quantum mechanics seem to be far away from any sense experiences. Penrose[4] 2016 has written the excellent book "FASHION, FAITH and FANTASY". On page 216 he writes:

> *Can* **fantasy** *have any genuine role to play in our basic physical understanding? Surely this is the very antithesis of what science is about, and should have no place in honest scientific discourse. However, it seems that this question cannot be dismissed as easily as might have been imagined, and there is much in the working of nature that appears fantastical, according to the conclusions that rational scientific thought appear to have led us to when addressing sound observational findings. As we have seen, particularly in the previous chapter, the world actually does conspire to behave in a most fantastical way when examined at a tiny level at which quantum phenomena hold sway. A single material object can occupy several locations at the same time and like some vampire of fiction (able, at will, to transform between a bat and a man) can behave as a wave or as a particle seemingly as it chooses, its behavior being governed by mysterious numbers involving the "imaginary" square root of -1.* Penrose 2016

Moreover, he said in an interview:

> *Physics is wrong, from string theory to quantum mechanics.* Penrose, 2009, DISCOVER

Weinberg[5] 2017 writes in a readable article about quantum mechanics, in particular, about the measurement problem:

> *Even so, I'm not as sure as I once was about the future of quantum mechanics. It is a bad sign that those physicists today who are*

---

[2]Einstein [1936]
[3]Quoted by Aage Peterson, Bulletin Atomic Scientists, 1963, Vol. 19, Issue 7, p. 12
[4]Penrose [2016, p.216]
[5]Weinberg [2017]

> *most comfortable with quantum mechanics do not agree with one another about what it all means. The dispute arises chiefly regarding the nature of measurement in quantum mechanics.* Weinberg 2017

Hossenfelder[6] writes in her recently published book "Lost in Math. How Beauty Leads Physics Astray":

> *Quantum mechanics is spectacularly successful. It explains the atomic world and the subatomic world with the highest precision. We've tested it upside-down and inside-out, and found nothing wrong with it. Quantum mechanics has been right, right, and right again. But despite this, or maybe because of this, nobody likes it. We've just gotten used to it.*
>
> *In a 2015 Nature Physics review, Sandu Popescu calls the axioms of quantum mechanics "very mathematical", "physically obscure", and "far less natural, intuitive and physical than those of other theories". He expresses a common sentiment. Seth Lloyd, renowned for his work in quantum computing, agrees that "quantum mechanics is just counterintuitive". And Steven Weinberg, in his lectures in quantum mechanics, warns the reader that "the ideas of quantum mechanics present a profound departure from ordinary human intuition".* Hossenfelder 2018

This small selection of statements of famous scientists are disillusioning. It is somewhat curious that physicists do not agree what quantum theory tells us, even after a century of discussions. There are a number of very good books that have a critical attitude, not just writing about what we know, but about what we do not know. A recommendable, critical, recently published book is written by Cham and Whiteson[7] with the telling title "We Have no Idea, A Guide to the Unknown Universe".

In these notes, we argue that quantum mechanics can be developed and formulated close to our sense experience, not counterintuitive, but very natural. Looking more deeply into experimental results, it seems that several paradoxes and riddles can be avoided, and the well-known measurement problem can be explained, when we understand two concepts, namely probability and time. We want to show that modifications of both concepts are useful for a better understanding. In our opinion, quantum mechanics is simply a probability theory about the reality which distinguishes between possibilities, internal possibilities and outcomes. It is a theory characterizing the future and telling us exactly what one should expect. Moreover, our approach contains a "single-world interpretation" that avoids many well-known paradoxes and interpretations such as "many worlds" or "many minds". Nevertheless, the presented concepts should be appropriate for teaching engineers.

The following notes are partially very different from what can be read in most textbooks. We try to avoid magical descriptions, and instead try to

[6] Hossenfelder [2018, p.119]
[7] Cham, Whiteson [2017]

maintain a more critical attitude. **A major goal of these notes is to stimulate students not to believe, but to ask. In particular, we do not present or repeat one of the current and widely accepted physical opinions.**

> *If a man will begin with certainties, he shall end in doubts; But if he will be content to begin with doubts, he shall end in certainties.*
> Francis Bacon, Advancement of Learning.

The strange paradox that "a single material object can occupy several locations at the same time" contradicts deeply our experience, was never observed, and was never measured. In 2017, I published my lecture notes "Quantum Information Theory for Engineers: An Interpretative Approach“, in the following shortly denoted by QUITE[8]. My aim was to be as close to sense experiences as possible. In particular, I argued that in quantum mechanics it is not necessary to believe that a single object can occupy several locations at the same time, provided we accept our daily observation that time is partitioned into past, present, and future[9]. Moreover, complex numbers, under mild conditions, turn out to be the maximal field of numbers, according to a Theorem of Hurwitz. They can be visualized as simple arrows in the plane and are very natural. It would be a surprise, if complex numbers would not be fundamental in physics [10]. The concepts entanglement [11], Heisenberg's uncertainty principle [12], and the theory of special relativity[13] are considered from a new perspective, by far not restricted to microscopic systems[14].

This article is a supplement to QUITE. One central point of view developed here is as follows: quantum mechanics is a fundamental probability theory for calculating numerical probabilities by generalizing Laplace's rules to complex numbers, hence really simple. The quantum calculus turns out to be the universal tool for computing the probabilities of outcomes or elementary events via probability amplitudes for possibilities and internal possibilities.

In QUITE, a basic knowledge of probability was assumed. This supplement tries to give some more insight into probabilistic concepts. One major goal is to build a bridge to the famous paper of Feynman[15] about non-relativistic quantum theory, written in 1948. The knowledge of several parts of QUITE is advantageous. However, we have referenced these parts in most cases.

This presentation is hopefully suitable for students studying engineering, but perhaps also for people interested in the philosophy of physics. It is written in the form of lecture notes. Therefore, many repetitions occur, as is the case when giving a lecture. It has the advantage that sections can be read partly independent of each other.

---

[8] Jansson [2017]
[9] Jansson [2017, Section 4.3]
[10] Jansson [2017, Section 2.2]
[11] Jansson [2017, Section 4.13]
[12] Jansson [2017, Section 4.17]
[13] Jansson [2017, Section 4.14]
[14] Jansson [2017, Section 4.11]
[15] Feynman [1948]

A short remark to the title of this supplement. "Free climbing" means that a climber on a rock uses only his hands, feet, and body to make upward progress, supporting his body in the vertical world. In other words, a free climber doesn't use any further technical support, except his sense experience together with his body. Similarly, these notes don't start with any of the various technical quantum postulates. We start with our sense experiences that can be used to reconstruct the technical properties of quantum theory and of the theory of relativity, both theories known as the fundamental theories in physics. However, this supplement does not match the actual consensus, although it generates the same mathematical formalism up to quantum field theories.

In the following we consider and discuss several experiments. The notion "experiment" has to be understood in its broadest sense. For defining probability we need an experimental situation, where we generally assume:

Postulate: The possible results of an experiment are mutually exclusive events. In other words, the experimental results form empirically decidable *alternatives*, which we call *outcomes* or *elementary events*. We can always distinguish between mutually exclusive events. They either happen or do not happen. But two or more elementary events cannot happen simultaneously. The set of outcomes forms the *sample space*. These are the fundamental assumptions in probability theory.

Obviously, this assumption contradicts the widely accepted, but never measured, opinion that "a single material object can occupy several locations at the same time". There are very many books on probability theory. In any case, we recommend the Handbook of Probability[16], and the many references therein.

The remainder of this supplement is organized as follows.

After the introduction, in the second chapter we consider in Section 2.1 Laplace's[17] definition of probability and its basic rules which were the standard for a long time. The Monty Hall problem, discussed in Section 2.2, emphasizes the importance of the exact knowledge of the sample space when solving probabilistic problems. Then many erroneous conclusions can be avoided. In Sections 2.3 and 2.4 two paradoxes are considered that violate the principle of indifference. These paradoxes suggest that the knowledge of the sample space alone is not sufficient for obtaining numerical probabilities. Further details of the experimental set up are required. Then, the basic axioms of classical probability and its relationship to relative frequencies are presented in Sections 2.5 and 2.6. In particular, the fundamental add-and-multiply rule, meaning that "probabilities for disjoint events are added, and probabilities for independent events are multiplied", is considered.

In the third chapter of these notes we describe how classical and quantum probability can be unified. In Section 3.1 we replace the concept of an external time parameter by the *trinity* future, present and past and show its

[16] Rudas [2008]

[17] Laplace [1814]

consequences. We discuss the differences and relationships between possibilities, outcomes, and facts. In Section 3.2, the *superposition principle* is shown as the unspectacular property of expressing possibilities of one machine in terms of the possibilities of other machines. This is a macroscopic interpretation of this fundamental quantum principle. Moreover, it turns out that quantum theory and classical probability theory are not different probability theories, but complement one another. It follows that quantum mechanics is a probability theory calculating probabilities for outcomes via probability amplitudes for possibilities. In the following Sections 3.3 and 3.4 we introduce three mathematical equivalent representations of possibilities and outcomes: the number, register, and vector representation. Then, in the following sections, we apply our concept to slit experiments, Hardy's paradox and the Frauchiger Renner paradox.

In the fourth chapter, we consider a reformulation of Feynman's famous approach to quantum theory, where we use the notions of the previous sections. In particular, we explain how the fundamental add-and-multiply rule together with the concept of distinguishing between possibilities, internal possibilities, and outcomes leads to a unification of classical probability theory and quantum theory. We show that the Dirac-Feynman rules are straightforward generalizations of Laplace's rules. In Sections 4.1 and 4.2 this probability theory is applied to space-time paths leading to Feynman's path integral. A sketch of the derivation of Schrödinger's wave equation, and thus proving that Feynman's formulation implies the ordinary formulation of quantum mechanics, is presented in Section 4.3. Hamiltonian mechanics is a mathematical formalism that provides a deeper understanding of classical mechanics and of quantum mechanics. In Section 4.4, we show the relationship between the Lagrangian and the Hamiltonian, thus leading to Feynman's path integral formulated in terms of the Hamiltonian. Finally, we show how our probability concept can be useful applied to quantum electrodynamics, our best physical theory explaining biology and chemistry.

In the fifth chapter of these notes we describe several aspects of measurements in quantum theory, including causality. The latter, although daily experienced, is rejected by some physicists.

Finally, some conclusions are given, and an appendix is attached containing "Keep in minds".

Feedback This text is free to download from the internet.

- `http://www.ti3.tuhh.de/jansson/`.

I am deeply grateful for corrections, comments, and suggestions:

- `jansson@tuhh.de`.

Acknowledgements I wish to thank Ulrike Schneider for her assistance in preparing this lecture notes, including graphics and tables. I wish to thank

Kai Torben Ohlhus and David Sills for their critical reading of the manuscript, their feedback, and their suggestions.

Hamburg, Germany, June 2019

Christian Jansson

2 Classical Probability

We consider some basic classical probability concepts. Moreover, we show by means of some paradoxes that probability is not self-evident.

2.1 Laplace Experiments

For a long time Laplace's[18] definition of probability was the standard: the probability of an event is the number of outcomes in favor of this event, divided by the number of all outcomes. Laplace begins with a series of principles of probability, including the classical statement:

> *The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible.*
> Laplace 1814

Using the common notion that the set of all possible outcomes is called *sample space* Ω, that any subset A of the sample space is called *event*, and that any element $\omega \in \Omega$ is called *elementary event* or *outcome*, Laplace has postulated:

- (*Unity outcome)*: If there are several outcomes all contributing equally, and it is agreed that neither seems favored over the other, all outcomes should be equally likely assigned with the unit 1.

- (*Addition rule)* The probability of an event is obtained by summing up over all outcomes contained in this event, where each term in the sum is equal to 1, and then by dividing by a normalizing constant, namely the number of all possible outcomes of the sample space. In other words, probability is the ratio of the favored elementary events to the total possible elementary events.

The first postulate that all outcomes have an equal probability, provided there is no known reason for treating certain outcomes differently, is also called the *principle of indifference*. For a large number of situations - fair coin or die toss and so forth - there is no preference or dependency between outcomes, and it is natural to assume that each outcome in the sample space is equally likely to occur.

The second postulate implies the generalized addition rule:

- (*Generalized addition rule*): For pairwise disjoint events $A_1, A_2, ..., A_l \subseteq \Omega$ it holds true:

$$\text{Prob}(A_1 \cup A_2 \cup ... \cup A_l) = \text{Prob}(A_1) + \text{Prob}(A_2) + ... + \text{Prob}(A_l). \quad (1)$$

[18] Laplace [1814]

Pairwise disjoint events represent *mutually exclusive alternatives*, that is, they either happen or do not happen, but two or more disjoint events cannot happen simultaneously.

Laplace has also considered how to calculate the probability of events or experiments that can be broken down into a series of steps happening independently. He formulated that for independent events the probability of the occurrence of all is the product of the probability of each. This can be deduced from the fundamental multiplication rule that forms the foundation for solving counting problems:

- (*Multiplication rule*) If we perform a sequence of experiments, say N, then the general principle of counting all outcomes is as follows: if the first experiment results in n_1 possible outcomes, and if for each of these n_1 outcomes there are n_2 possible outcomes of the second experiment, and if for each of these $n_1 \cdot n_2$ outcomes there are n_3 possible outcomes of the third experiment, ... then there there is a total of $n_1 \cdot n_2 \cdot ... \cdot n_N$ outcomes of the sequence of experiments. In other words, the sample space of this sequence has $n_1 \cdot n_2 \cdot ... \cdot n_N$ elements.

Events are called *independent*, if the entering of one event does not influence the probabilities of the other events. Therefore, if in a sequence of N independent experiments the events $A_j, j = 1, ..., N$, have m_j possible outcomes, then the number of all possible outcomes is the product $m_1 \cdot ... \cdot m_N$, and the second principle implies the probability

$$\text{Prob}(A_1 \cap A_2 \cap ... \cap A_N) = \frac{m_1 \cdot m_2 \cdot ... \cdot m_N}{n_1 \cdot n_2 \cdot ... \cdot n_N} = \text{Prob}(A_1) \cdot \text{Prob}(A_2) \cdot ... \cdot \text{Prob}(A_N). \tag{2}$$

These principles apply to a large number of experiments provided the sample space is finite. This requires, however, to define precisely the conditions of an experiment, that is, to make a list of all possible outcomes, and to assure that all outcomes are equally likely.

Keep in mind: Calculate the probability for the outcomes in Laplace experiments by using the *multiply-and-add rule*, that is, the probabilities for disjoint events are added, and the probabilities for independent events are multiplied. This rule is universal, since it applies also to classical probability as formulated by Kolmogorov, and to quantum probability.

There are numerous experiments, however, that cannot be solved with these rules. A simple example is to toss a pushpin. The probability that it ends up on its head is in general not equal to the probability that the pin ends up on the other side. These probabilities depend on the specific geometry of the pin. Or, think of throwing darts on a disc, which is partitioned into different sections. If we never miss the disc and do not aim to hit a special section, then the probability of hitting some section should be the area of this section divided by the area of the disc.

Before we describe other concepts of probability, we proceed with some seemingly paradoxical experiments, see also the Handbook of Probability[19].

[19] Rudas [2008]

2.2 Monty Hall Problem

The *Monty Hall problem* is a statistical puzzle named after Monty Hall, the host of the television game show "Let's Make a Deal". This puzzle was posed (and solved) by Steve Selvin in 1975, and became famous by Marilyn vos Savant's "Ask Marilyn" column in *Parade* magazine in 1990. It is as follows: In a game show you are given the choice of three doors, where behind one of the doors is a car, and behind the other two doors are goats. If you choose a door, the host, who knows what is behind the doors, opens another door with a goat behind. Then he asks whether you want to switch your chosen door or you want to stay. Hence, the question arises, whether it is better to switch your choice?

Most people argued that switching the door is not necessary, since there are two unopened doors with one goat and one car behind. Thus, using the *principle of indifference*, we get a 50/50 chance. It turns out, however, that switching has a 2/3 chance of winning the car, while staying at the chosen door has only a 1/3 chance, as correctly stated by Marilyn vos Savant.

The Monty Hall problem has attracted a lot of attention. About 10,000 readers of the magazine, including many PhDs, wrote to the magazine claiming that vos Savant is wrong. Even the famous mathematician Paul Erdös remained unconvinced until a computer simulation was given to him[20] . Surprisingly, this problem seems to have been so interesting that a book about the letters from the readers of the magazine was written[21].

There are many solutions for this problem, including approaches using conditional probabilities, Bayes Theorem, and several other ideas. But the key insight can be obtained when looking carefully at the sample space of this problem. Almost everyone knows that defining the outcomes or elementary events forms the basis for solving statistical problems.

Let's do this. In fact, we have two problems. First, the decision is to stay at the chosen door. Since the numbering of the doors does not matter, we suppose that the chosen door is door 1. Then we obtain exactly three outcomes displayed in Table 1. All outcomes are equally likely, and we can apply the principle of indifference and Laplace's rules. Only in one case you win the car. Hence, staying at the chosen door has only a 1/3 chance of winning the car.

Behind door 1	Behind door 2	Behind door 3	Stay at door 1
Goat	Goat	Car	Gets goat
Goat	Car	Goat	Gets goat
Car	Goat	Goat	Gets car

Table 1: The outcomes for the Monty Hall problem if you stay.

Secondly, the decision is to switch the chosen door 1. We obtain three outcomes as before, displayed in Table 2. Now in two cases you win the car. Hence, switching the chosen door has a 2/3 chance to win the car. Summariz-

[20]Vazsonyi [1999]
[21]Granberg, Brown [1995]

ing, the solution is almost trivial, when looking carefully at the sample space and identifying the problem as a simple Laplace experiment.

Behind door 1	Behind door 2	Behind door 3	switch
Goat	Goat	Car	Gets car
Goat	Car	Goat	Gets car
Car	Goat	Goat	Gets goat

Table 2: The outcomes for the Monty Hall problem if you switch your choice.

Finally, let us change the problem such that the host opens the door at random rather than always revealing a goat. Then the probability changes to 1/2, because 1/3 of the time he opens the door with the car behind, therefore ending the game.

Keep in mind: When solving probabilistic problems it is necessary to know precisely the sample space. Then many erroneous conclusions can be avoided, as the letters to Marilyn vos Savant demonstrate.

2.3 Bertrand's Chord Paradox

The Monty Hall problem is a so-called weak paradox, that is, a problem demonstrating the weakness of understanding probability. But it can be solved with a little thought. There exist various strong paradoxes that apparently seem to falsify classical probability theory and pose challenges and deep problems to the classical theory[22].

Bertrand designed one type of well-known paradoxes. His aim was to argue that the *principle of indifference* is not applicable to experiments with infinitely many outcomes, and as proof, he offered some examples leading to contradictions. This includes his famous chord paradox:

> *We trace at* **random** *a chord in a circle. What is the probability that it would be smaller than the side of the inscribed equilateral triangle?* Bertrand[23]

For solving this problem we can use the principle of indifference in three different ways. A chord intersecting the circle is uniquely defined by its two points intersecting the circumference. Hence, tracing out at random a chord can be done by generating at random these two points. We consider three possibilities for obtaining a probability:

- 1) Take one of the two points as the vertex A of the inscribed equilateral triangle ABC. Then the chord is longer than the side of the triangle, if it lies within the angle at the vertex A. This is true for one-third of the chords, since the angle is 60 degree compared with 180 degree at A. Hence, the probability of getting a longer chord is one-third, and the probability of a smaller chord is two-thirds.

- 2) But we can argue also in another manner. Given the chord, let the inscribed triangle be defined such that the side BC is parallel to the chord. Then the midpoints of the chord and the side lie on the same radius perpendicular to the chord and the side. Then the chord is smaller if it intersects the outer half of the radius perpendicular to them, so that their midpoint is outside the triangle. Hence, the probability is one-half.

- 3) Finally, a chord is longer than the side of the triangle, if its midpoint falls within a circle inscribed within the inscribed equilateral triangle. This inner circle has a radius one-half. Hence, its area is one-quarter of the outer circle, leading to probability one-quarter. Thus, the probability of chords being shorter is three-quarter.

We used in all cases the principle of indifference, frequently supposed to suffice for solving probability problems. By definition, probabilities have unique solutions, because they are described as a single function from the events of interest into the interval $[0, 1]$. But now, three different ways of applying this

[22] [Chapter 3]Rudas [2008]
[23] Bertrand [1889, pp. 4]

principle result in different probabilities for the same event. Bertrand concludes that this example undermines the principle of indifference.

In the "Handbook of Probability" it is stated that Bertrand's paradoxes are unresolved, thus threatening our confidence when applying probability theory to infinite sets[24]. But is this really a paradox? Perhaps not. You can write three programs for the three cases. Then these programs are different and produce approximately the probabilities above, provided your programs are written correctly. The reason is that even a precisely defined sample space does not necessarily imply correct probabilities. The experimental set up must be incorporated appropriately, as in the programs. In fact, you might realize these three cases also physically.

Keep in mind: When solving probabilistic problems, a precisely defined sample space may be not sufficient. The "principle of indifference" may be violated. For obtaining numerical probabilities, the process or program how the outcomes of the sample space are constructed may be necessary.

[24][pp. 54]Rudas [2008]

2.4 Bertrand's Cube Paradox

Here, we consider a paradox of Bertrand in the form adapted from van Fraassen 1989, see also the "Stanford Encyclopedia of Philosophy: Interpretations of Probability" and "Philosophies of the Sciences: A Guide[25]". We quote Lyon, see the chapter in the latter book.

> *Consider a factory that produces cubic boxes with edge lengths anywhere between (but not including) 0 and 1 meter, and consider two possible events: (a) the next box has an edge length between 0 and 1/2 meters or (b) it has an edge length between 1/2 and 1 meters. Given these considerations, there is no reason to think either (a) or (b) is more likely than the other, so by the Principle of Indifference we ought to assign them equal probability: 1/2 each. Now consider the following four events: (i) the next box has a face area between 0 and 1/4 square meters; (ii) it has a face area between 1/4 and 1/2 square meters; (iii) it has a face area between 1/2 and 3/4 square meters; or (iv) it has a face area between 3/4 and 1 square meters. It seems we have no reason to suppose any of these four events to be more probable than any other, so by the Principle of Indifference we ought to assign them all equal probability: 1/4 each. But this is in conflict with our earlier assignment, for (a) and (i) are different descriptions of the same event (a length of 1/2 meters corresponds to an area of 1/4 square meters). So the probability assignment that the Principle of Indifference tells us to assign depends on how we describe the box factory: we get one assignment for the "side length" description, and another for the "face area" description.*
>
> *There have been several attempts to save the classical interpretation and the Principle of Indifference from paradoxes like the one above, but many authors consider the paradoxes to be decisive. See Keynes [1921]3 and van Fraassen [1989]4 for a detailed discussion of the various paradoxes, and see Jaynes [1973]5, Marinoff [1994]6, and Mikkelson [2004]7 for a defense of the principle. Also see Shackel [2007]8 for a contemporary overview of the debate. The existence of paradoxes like the one above were one source of motivation for many authors to abandon the classical interpretation and adopt the frequency interpretation of probability.* [...]
>
> *Ask any random scientist or mathematician what the definition of probability is and they will probably respond to you with an incredulous stare or, after they have regained their composure, with some version of the frequency interpretation. The frequency interpretation says that the probability of an outcome is the number of experiments in which the outcome occurs divided by the number of experiments performed (where the notion of an "experimen" is understood very broadly). This interpretation has the advantage that*

[25] Allhoff [2010]

> *it makes probability empirically respectable, for it is very easy to measure probabilities: we just go out into the world and measure frequencies.* Lyon 2010

What then is the probability of the previous event in question?

A little bit more insight can be obtained when we use instead of continuous intervals a discrete version of the paradox. We suppose that the factory produces only boxes that have edge lengths $1/4,\ 1/2,\ 3/4,\ 1$. Thus, their face areas are $1/16$,$1/4,\ 9/16,\ 1$, and in both cases the principle of indifference tells us that the probability is $1/2$ for the event that the next box has an edge length between 0 and $1/2$ meters or a face area between 0 and $1/4$ square meters. That look's nice. Conversely we suppose now that the factory produces only boxes that have face areas $1/4,\ 1/2,\ 3/4,\ 1$. Thus, their edge lengths are $1/2,\ 1/\sqrt{2},\ \sqrt{3}/2,\ 1$, and in both cases we obtain, by using the principle of indifference, probability $1/4$. In summary, for the discrete version of this paradox, the probabilities for both events are identical. One might say that the paradox vanishes in the discrete case, but occurs when we pass to the continuous case.

Similarly, as in the previous chord paradox, it is a different situation whether the factory produces the boxes by firstly choosing the edge length, or by firstly choosing the face area. This can be verified by writing a program. The mathematical reason behind is the fact that a uniform distribution entails a non-uniform distribution when transforming nonlinear. This can be seen immediately from the discrete version above. There, we have squared the edge length in order to obtain the face area. Thus, a uniform distribution remains no longer uniform under quadratic transformations, and the principle of indifference does not apply in the infinite case.

Keep in mind: In general, under nonlinear transformations the type of distributions changes, and the principle of indifference does not apply.

2.5 Kolmogorov's Axiomatization

We have seen that the assumption of a finite sample space is sometimes unsatisfactory, remarked already by Bertrand 1889. Rather late in 1933, Kolmogorov presented a mathematical theory of probability in terms of some axioms which have become orthodoxy. Actually, it is a measure theory. Measures generalize volumes and are used in several non-probabilistic applications, for instance, Lebesgue measures and integration. The axioms in measure theory, however, do not calculate probabilities for outcomes like Laplace did.

In the following, the *sample space* Ω of outcomes may be finite or infinite. A *field* is a set of subsets of Ω that contains the sample space itself, and is closed under the countable set operations union, intersection, and complement. Thus, for countable subsets $A_1, A_2, ...$ the set $A_1 \cap A_2 \cap ...$, the set $A_1 \cup A_2 \cup ...$, and the set $A_m - A_n$ are elements of the field. The elements of the field are called *events*. The outcomes are called *elementary events*.

The basic axiom is to assign a mapping Prob, called *probability function*, from the field of events

$$A \to \text{Prob}(A) \tag{3}$$

into the set of real numbers that satisfies

$$0 \leq \text{Prob}(A) \leq 1, \quad \text{Prob}(\Omega) = 1, \tag{4}$$

and moreover for any countable set of disjoint events A_m the equation

$$\text{Prob}\left(\bigcup_{m=1}^{\infty} A_m\right) = \sum_{m=1}^{\infty} \text{Prob}(A_m). \tag{5}$$

must be fulfilled.

The axiom (4) is important when performing the same experiment several times. Otherwise, we cannot hope that the relative frequencies of an event A approaches Prob(A). The *relative frequency* is the number of times the event A occurred in a series of executions of an experiment divided by the number of executions, thus is bounded between 0 and 1.

Two events A and B are called *independent*, if both have no influence on each other. For instance, if we toss a coin twice, and we know the outcome A of the first toss, then this has no influence on the result B of the second toss. In accordance with Laplace experiments, the probabilities for independent events are multiplied, that is,

$$\text{Prob}(A \cap B) = \text{Prob}(A)\text{Prob}(B). \tag{6}$$

In summary, the probabilities of disjoint events are added, and the probabilities of independent events are multiplied. This is the well-known *multiply-and-add rule* **which holds valid already for Laplace experiments, but now for non-negative real numbers.**

From these axioms one can deduce the well-known and useful rules for calculating probabilities, provided the probabilities for the outcomes are given, or some distribution. We assume only a rudimentary knowledge of the most important probabilistic rules.

The mathematics of Kolmogov's probability theory is well understood, but its interpretation is controversial. A nice survey about various other interpretations is written by Hajek[26]. In particular, there are also many non-Kolmogorovian theories of probability. Usually, a probability is a single number. But there are approaches that use interval-valued probabilities, or offer axioms for "upper" and "lower" probabilities. Even some scientists drop the normalization assumption altogether, allowing probabilities to attain the value ∞. We mention this for the interested reader, but we don't go into details.

[26] Hajek [2001]

2.6 Probability and Relative Frequency

Above, we have already mentioned:

> *Ask any random scientist or mathematician what the definition of probability is and they will probably respond to you with an incredulous stare or, after they have regained their composure, with some version of the frequency interpretation.* Lyon 2010

In experiments we can always distinguish between mutually exclusive *outcomes*. They either happen or do not happen, but two or more outcomes cannot happen simultaneously.

This assumption seems to be trivial. It is accepted everywhere when solving classical probabilistic problems. It seems to be accepted in quantum theory, since all (well-working) detectors display only one outcome. However, remembering the quotation of Penrose in the introduction, a single material object can occupy many different places, although this was never measured. This is a strange paradox in quantum mechanics. In QUITE, however, it is argued that a single material object occupies exactly one location. Hence, in the following we always use our assumption of mutually exclusive outcomes.

Let us perform the same experiment N times, and suppose that an outcome, say k, occurred n_k times in this series. Then we call the ratio

$$f_k = \frac{n_k}{N} \tag{7}$$

the *relative frequency* corresponding to outcome k. This ratio is a rational number between 0 and 1. Intuitively, we expect that the relative frequency is a number close to a probability of this outcome, at least when we repeat this experiment a lot of times.

When tossing a fair coin, for example, it seems plausible that the relative frequency is close to 1/2 for the outcome "heads". Also for throwing a fair die f_k should be close to 1/6 for all outcomes $k = 1, 2, 3, 4, 5, 6$. Now it seems to be natural to define a probability Prob(k) for outcome k as the limit of f_k for $N \to \infty$. This point of view is the *frequentist definition* of probability, providing an operational definition.

It is perhaps the most widespread imagination of probability, although there are several disadvantages. Experiments can be repeated only a finite number of times, even if they could be performed infinitely often in principle. Moreover, this definition could be misleading, since even a fair coin might land heads 99 out of 100 times. However, there is a probabilistic model supporting this operational definition. It is the *law of large numbers*, stating that for an experiment, performed many times, the relative frequency tend to be close to the theoretical probability of this model. But this definition does not work in general, since many experiments are unrepeatable, such as elections or events in sports. This problem of the so-called "single case" is striking, since the frequentist definition does not apply. For the moment, however, we can take this definition as an operational point of view of probability.

3 Unification: Classical and Quantum Probability

So far we have discussed classical probabilities and some paradoxes. We have not dealt with quantum probability, probability amplitudes, and interference yet. Such phenomena have been investigated 1948 by Feynman[27]. He described a third formulation of quantum theory, interpretative completely different to previous insights. His formulation can be viewed as a non-classical probability theory with complex numbers, not with real positive ones. He proved this formulation to be mathematically equivalent to Schrödingers wave theory and the matrix algebra of Heisenberg. What is the value of mathematical equivalent descriptions? Well, he could describe old things from a new point of view, and could offer distinct advantages of his formulation. In particular, he discussed from his perspective the superposition in quantum mechanics, wave equations, operator algebra, the relationship to a large class of action functionals, commutation relations, Newton's mechanics, statistical mechanics, and ideas of his extension to quantum electrodynamics. It is a wonderful written paper containing all fundamental concepts of quantum theory on a few pages. It is simply great, and I recommend anyone interested in quantum theory to read it.

Feynman's formulation is based on spacetime. He used a typical physical language in his article. He argued with measurements, and thus came to wave-particle duality and the related *measurement problem* in quantum theory, namely the well-known problem whether and how the wave function collapses to a certain event or outcome. This questioning has pushed one of the most challenging and partially nebulous debates about reality and quantum theory.

Our approach replaces spacetime by a trinity of time, namely future, present, and past. Moreover, we use a set-theoretical language that is typical in classical probability theory, as described in the previous sections. This offers a completely distinct interpretation without paradoxes and riddles, and many things become much clearer. In particular, we consider the difference between outcomes, possibilities, and internal possibilities. It turns out that these quantities depend on time. We obtain a probability theory that unifies the *Dirac-Feynman rules* with classical probability theory. We show that the Dirac-Feynman rules are straightforward generalizations of Laplace's rules. We discuss polarization and slit experiments, as well as Hardy's paradox, a spectacular experimental setup where simple logical arguments about its physical constraints lead to a surprising contradiction. However, we can explain these experiments in a rather simple manner, without any strange arguments.

[27]Feynman [1948]

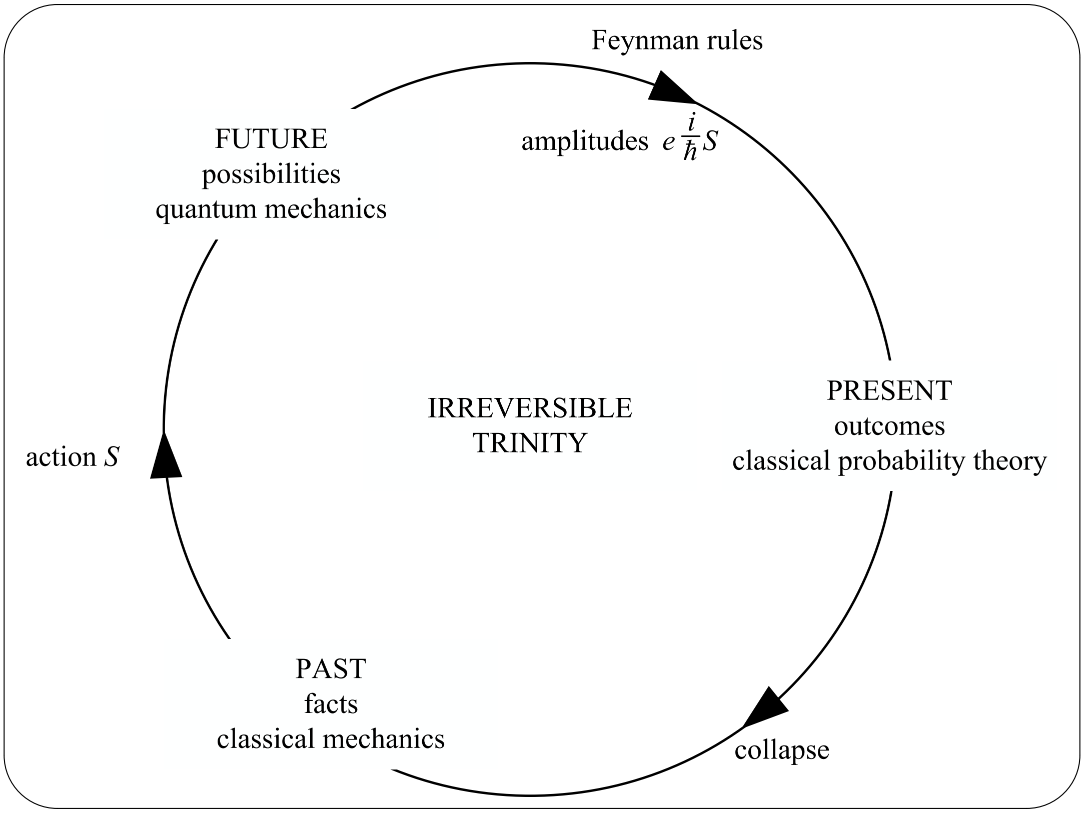


Figure 1: Trinity of time.

3.1 Trinity of Time

Quantum theory can be understood rather easily when we replace the concept of an external time parameter t, generally used in physics, by the trinity future, present, and past, see Figure 1. It is very close to our sense experiences. In this section, we present a short and rough overview. Precise definitions are given afterwards. More details and several applications are considered and discussed in QUITE.

There, we have interpreted quantum mechanics as a theory of probabilistic predictions that characterize the future only. The *future* is timeless, nothing happens, and it might be best described by the phrase "What might happen, when nothing happens?". In other words, quantum mechanics has to be understood prognostically. It is a probability theory that assigns to mutually exclusive *possibilities* complex numbers, the so-called *probability amplitudes*. We look in the following at three types of experiments: throwing a die, the slit experiment, and the polarization of photons.

When throwing a fair die, we obtain six mutually exclusive possibilities $k = 1, 2, 3, 4, 5, 6$ with probability amplitudes $1/\sqrt{6}$. Squaring gives the probabilities $1/6$.

In a double-slit experiment, see Figure 2, the paths from a fixed initial point s via any slit to any final point at the screen, here defined as a position detector d_m, describe the possibilities. They are allocated with complex probability

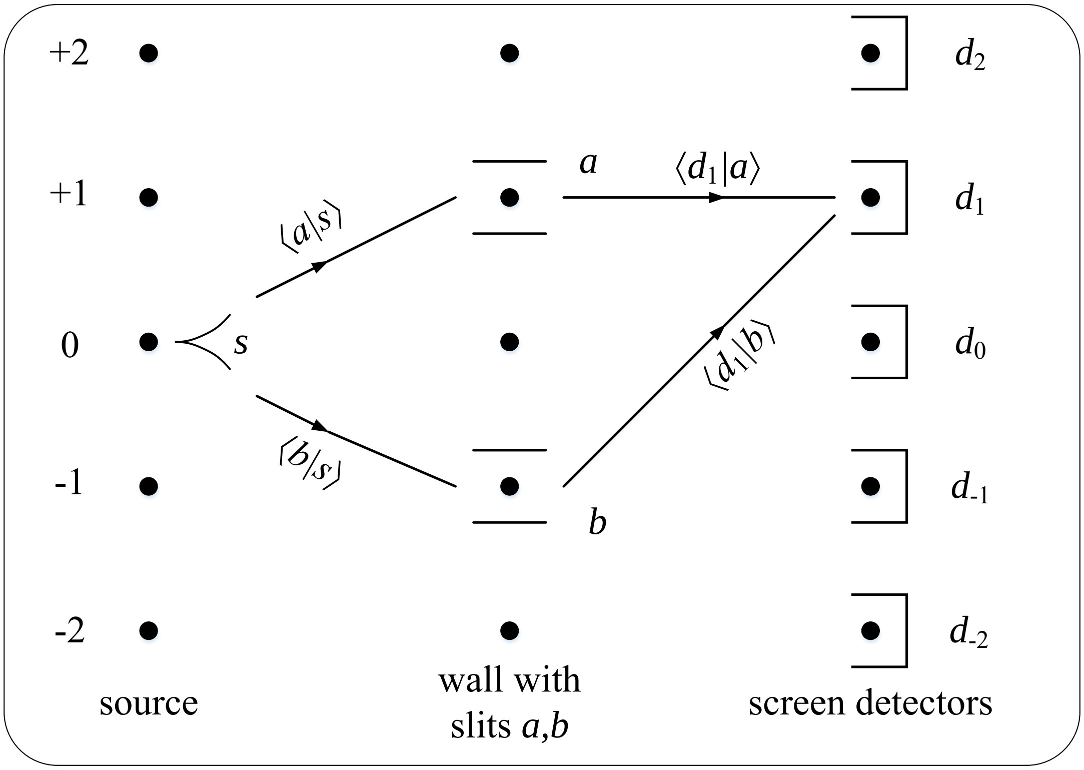


Figure 2: The double-slit experiment described for a discrete spacetime. The particle leaves source s, passes one of the two slits a or b, and is finally detected in d_1.

amplitudes[28].

Finally, we consider the polarization experiment in Figure 3. The mutually exclusive possibilities in a future execution are:

- (1) The photon is absorbed by the first polarizer.
- (2) The photon passes the first polarizer, then moves on the upper beam between the birefringent plates, and finally is absorbed by the second polarizer.
- (3) The photon passes the first polarizer, then moves on the lower beam between the birefringent plates, and finally is absorbed by the second polarizer.
- (4) The photon passes the first polarizer, then moves on the upper beam between the birefringent plates, and finally passes the second polarizer, detected after that.
- (5) The photon passes the first polarizer, then moves on the lower beam between the birefringent plates, and finally passes the second polarizer, detected after that.

[28] Jansson [2017, Sections 2.6 and 2.7]

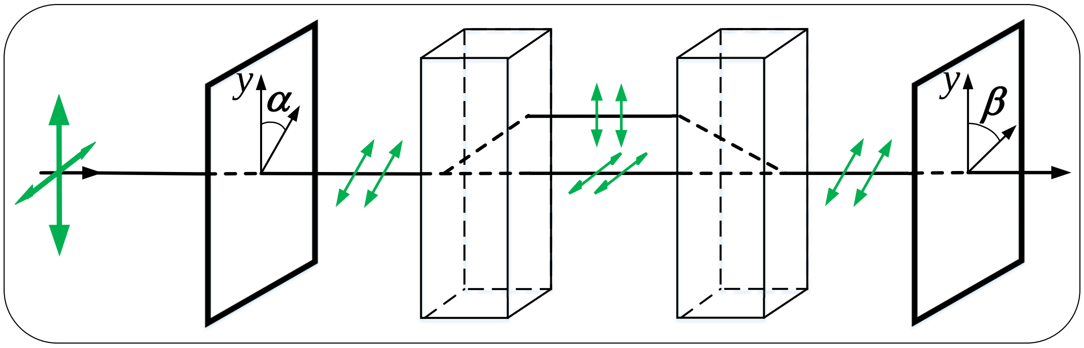


Figure 3: The first polarizer generates photons polarized at an angle α. The first birefringent plate splits into two beams of horizontally x-polarized and vertically y-polarized photons. These are recombined in a second birefringent plate which has an optical axis opposite to the first plate. According to the law of Malus the transition probability after the second polaroid is $\cos^2(\beta-\alpha)$.

So far to the prognostic future. In the present, experiments are performed. The present is characterized by classical random access. In the present momentary decisions take place. The possible results, expressed by the detectors, are called *outcomes* or *elementary event*. They define the *sample space*. In general, possibilities and outcomes differ. The outcomes are those possibilities that represent possible interactions with detectors or the environment, whereas the remaining ones are the internal alternatives which we call *internal elementary possibilities*. We call physical models *classical*, if the possibilities coincide with the outcomes, that is, internal elementary possibilities are not present.

When throwing a fair die, the table where the die is finally located acts as a detector. Possibilities and outcomes don't differ for this example; they are the numbers $k = 1, 2, 3, 4, 5, 6$. Hence, we have a classical model.

In the slit experiment without detectors at the slits, a particle follows exactly one path in the present, from the starting point to any position at the wall of detectors. The positions at the last wall form the outcomes. But there are many paths through the slits, describing internal possibilities, that lead to the same outcome. This is a non-classical model. However, if we position detectors at the slits, then we obtain a classical model.

The outcomes for the polarization experiment in Figure 3 are:

- (1) The photon is absorbed by the first polarizer.
- (2) The photon passes the first polarizer, then moves through the birefringent plates, and finally is absorbed by the second polarizer.
- (3) The photon passes the first polarizer, then moves through the birefringent plates, and finally moves through the second polarizer, detected after that.

Hence, five possibilities are reduced to three (detected) outcomes. It is a non-classical model. The possibilities, describing what happens between the birefringent plates, are internal, that is, they are not given to the environment.

In fact, this characterizes a fundamental difference between future and present. The property that there may be more possibilities than outcomes becomes incomprehensible when using spacetime only. Actually, models based solely on spacetime lead to statements like "a material objects occupies several locations at the same time“. The failing distinction between past, present, and future in physics, is the reason for many paradoxes in current quantum theory.

Deterministic models, like classical mechanics or electromagnetism, are described uniquely in terms of differential equations that don't allow alternative solutions. There is a unique outcome changing deterministically with time, yielding a classical model. Statistical mechanics is classical, since there are no internal elementary possibilities. All possibilities are outcomes. Quantum mechanics is non-classical, since outcomes can be reached via several internal elementary possibilities. Summarizing, we have precisely defined the notion "classical". In the literature, this notion is vague.

In statistical mechanics the concept *probability* is defined mathematically as a map from the set of all outcomes, namely the sample space, into the set of real numbers between zero and one. Since probabilities are non-negative numbers, cancellation or interference does not occur. It is a notion of the present, where probabilities determine which of the outcomes momentarily becomes a fact. A *probability amplitude* is defined as a map from the set of all possibilities into the set of complex numbers with magnitudes between zero and one. Squaring the magnitude of probability amplitudes gives the probabilities, according to *Born's rule*. Probability amplitudes are the quantities that can describe appropriately geometric details of the experimental set up. Since these are complex numbers, cancellation producing interference phenomena may occur.

In the *past*, one of the outcomes has become a *fact*. The past is deterministic, and classical mechanics can be viewed as a theory of the past. The concept of *relative frequencies* describe the outcomes or measured results of repeated experiments, and thus belongs to the *past*. Not surprisingly, the past serves to verify or falsify prognostic statements. But from the philosophical point of view, however, it is doubtful to define probabilities for events that didn't happen via concepts of the past.

It is important to notice that in our approach possibilities are properties of the machines that form the experimental set up, as seen above. Possibilities represent mutually exclusive *alternatives* in the sense that in a future experiment, a particle interacting with a machine, chooses exactly one of these alternatives, not two or more. For example, polarization is first and foremost a property of the optical apparatus, not of a photon itself. We can only say that a photon interacts in the present with a specific crystal or polarizer by choosing exactly one of its possibilities. A single material object doesn't occupy several locations at the same time. It chooses in the present exactly one location.

This trinity is close to experience. Learning would be impossible, if we don't distinguish between things that might happen and things that have happened. Time is one of the most discussed concepts in physics and philosophy. Here, we quote two prominent physicists. In the introduction we have already quoted Einstein's article[29] about physics and reality. He claimed at various pages closeness to sense experiences. However, when he considered the concept of time[30] he changed from an "experienced local time", which connects the temporal sequence of experiences, to an "objective time":

> *The concept of space is, it is true, useful, but not indispensable for geometry proper, i.e. for the formulation of rules about the relative positions of rigid bodies. In opposition to this, the concept of objective time, without which the formulation of the fundamentals of classical mechanics is impossible, is linked with the concept of the spatial continuum.*
>
> *The introduction of objective time involves two statements which are independent of each other.*
>
> *(1) The introduction of the objective local time by connecting the temporal sequence of experiences with the indications of a "clock" i.e. of a closed system with periodical occurrence.*
>
> *(2) The introduction of the notion of objective time for the happenings in the whole space, by which notion alone the idea of local time is enlarged to the idea of time in physics.*
>
> *[....]*
>
> *The illusion which prevailed prior to the enunciation of the theory of relativity - that, from the point of view of experience the meaning of simultaneity in relation to happenings distant in space and consequently that the meaning of time in physics is a priori clear, - this illusion had its origin in the fact that in our everyday experience, we can neglect the time of propagation of light. We are accustomed on this account to fail to differentiate between "simultaneously seen" and "simultaneously happening"; and, as a result the difference between time and local time fades away.*
>
> Einstein 1936

In physics, time t appears in almost all equations. Physicists think that these equations describe what happens in the next moment. Variables such as the position $x(t)$, the velocity $v(t)$, the momentum $p(t)$, the energy $E(t)$, and so on, are time-dependent. In the case of the harmonic oscillator, the well-known Euler-Lagrange equation takes the form of a differential equation

$$\frac{d}{dt}(m\dot{x}) - kx = 0. \tag{8}$$

The idea of equations without the variable time seems questionable at first, or even very strange. But after a while, you realize that the variable time is not necessary. We can establish timeless relationships between the other variables. For the harmonic oscillator, for instance, the *Hamiltonian*

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 \tag{9}$$

[29] Einstein [1936]

[30] Einstein [1936, pp. 357]

is the conserved total energy, that is, the sum of kinetic and potential energy. This equation describes the harmonic oscillator just as well without t, implicitly. It represents an ellipse in the phase space. Exactly the same situation can be found in the famous *Wheeler-de Witt equation*, a candidate for the solution of the well-known quantum gravitation problem. This equation contains no time parameter. The time-dependent equations don't describe what happens in the next moment, but describe deterministic quantities belonging to the past in its explicit form.

In QUITE, Sections 4.13 and 4.14, we gave several arguments to choose an Euclidean (3+3)-position-velocity space as a basis of physics, without any time parameter. The major reason was that we view quantum mechanics as a theory describing the future, and the future is timeless. But then spacetime vanishes, and there seems to be no theory of relativity. Nevertheless, it was shown to reproduce the mathematical formalism of special relativity by constructing clocks in the (3+3)-position-velocity space. In particular, we derived the key of relativity theory, namely the *Lorentz transform*, without any assumption about "propagation of light". Hence, both statements of Einstein above, "the concept of objective time, without which the formulation of the fundamentals of classical mechanics is impossible“ and "this illusion had its origin in the fact that in our everyday experience, we can neglect the time of propagation of light", can certainly be questioned. It was shown that the Euclidean position-velocity space[31], being close to our sense experiences, allows us to describe Hamilton's classical mechanics, the theory of special relativity, and a reasonable explanation of entanglement.

In summary, we have three models related to future (*quantum mechanics*), present (*statistical mechanics*), and past (*classical mechanics*). These are timeless theories in the sense that we don't need an external time parameter and spacetime. In classical mechanics, for instance, the time parameter t is a geometric parameter that serves to represent a classical solution in an explicit form. At a first glance, this trinity seems to create another time concept when the circle in Figure 1 rotates. However, this concept is completely different from other time ideas since it rotates the past into the future, the future into the present, and the present into the past. Moreover, it is completely different to the imagination of an arrow of time.

Von Weizsäcker[32] emphasizes at various places the fundamental difference between the "factual past“ and "possible future“. Using the language of temporal logic, he distinguished between "presentic, perfectic, and futuric statements“. However, he returned to spacetime by investigating the quantum theory of binary alternatives.

[31]Geometrically described by the isomorphic Lie algebras $so(4) \cong so(3) \times so(3) \cong su(2) \times su(2)$

[32]von Weizsäcker [1988], von Weizsäcker [1992], von Weizsäcker [2006]

Keep in mind: Bertrand's paradoxes have shown that the sample space is not sufficient for calculating probabilities. Further information about the experiment is necessary. This information depends on the geometry of the experimental set up. Moreover, we assume a time trinity that distinguishes between future possibilities, present random access of outcomes in terms of momentary decisions, and the facticity of the past in terms of facts. Facts are elements of the set of outcomes, the latter are contained in the set of all possibilities. Time trinity allows, in a very simple way, to describe precisely experiments. A *probability* is defined as a map from the set of all outcomes into the set of real numbers between zero and one, and is related to the present. A *probability amplitude* is defined as a map from the set of all possibilities, including *internal elementary possibilities*, into the set of complex numbers with magnitudes between zero and one, and is related to the future. Squaring the magnitude of probability amplitudes for outcomes gives the probabilities, according to Born's rule.

3.2 The Superposition of Probability Amplitudes

We consider an imaginary experiment consisting of three machines A, B, and C connected in series[33]. The machines can interact with a specific type of particles. Which type doesn't matter in the following. The machines are characterized by its elementary mutually exclusive *alternatives*, that is, the elementary *possibilities* $a \in A, b \in B, c \in C$. *Elementary* means that the possibilities cannot be further separated. *Mutually exclusive* means that the possibilities are non-overlapping and distinguishable. Moreover, we assume that, in the present, a particle can interact with a machine by choosing exactly one possibility, but two or more possibilities cannot be chosen simultaneously. Consequently, viewing space as a machine of positions, a single material object cannot occupy several locations simultaneously.

We assume that the set of elementary possibilities is countable. The elementary possibilities of the complete experiment ABC consist of all triples abc. It says that, in a future interaction of a particle with the experimental set up, the particle chooses possibility a, then b, and finally c. We call the set of all *elementary possibilities* abc the *possibility space* P of the experiment, that is,

$$P = \{abc : \ a \in A, b \in B, c \in C\}. \tag{10}$$

The experiment itself can be viewed as one single machine.

The field F_P is defined as the set of all subsets of the possibility space, that is,

$$F_p = \{abc, abC, aBc, Abc, aBC, AbC, ABc, P, \emptyset, \text{ where } a \in A, b \in B, c \in C\}. \tag{11}$$

There, the elementary possibilities abc, which we identify with $\{abc\}$, are the subsets consisting of one element. The other subsets are the *non-elementary possibilities* defined as

$$abC := \{abc : \ c \in C\}, \tag{12}$$
$$aBc := \{abc : \ b \in B\}, \tag{13}$$
$$Abc := \{abc : \ a \in A\}, \tag{14}$$
$$aBC := \{abc : \ b \in B \ c \in C\}, \tag{15}$$
$$AbC := \{abc : \ a \in A \ c \in C\}, \tag{16}$$
$$ABc := \{abc : \ a \in A \ b \in B\}, \tag{17}$$
$$ABC := P. \tag{18}$$

For instance, the possibility aBc means that, in a future interaction of a particle with the experimental set up ABC, the particle chooses the elementary possibility a, finally has chosen c, and further it must have chosen some intermediate, not further specified, elementary possibility b provided by machine B.

[33]See Section 2, Feynman [1948]

It may be that we are not interested in the possibilities of B. But it may also be that the interaction with B is not known and cannot be determined. In other words, it is not given outside to the environment. Then we say that the possibilities $b \in B$ are *internal*. It turns out that **the internal possibilities of an experimental set up must be explicitly defined. They are responsible for** *interference*. We speak of a *classical experiment*, if internal possibilities do not occur.

The double slit experiment, described in Figure 2, consists of three machines denoted by SWD. The first machine denotes the source S producing particles, the second machine W is the wall with two slits without detectors, say a and b, and the third machine D is the screen of position detectors d_m. Since there are no detectors at the slits, the possibilities of W, representing both slits, are internal. In the present, it is not given to the environment through which slit the particle passes, yielding a non-classical experiment. This experiment becomes classical, if we put detectors at the slits.

Notice, we consider future interactions that do not happen, but might happen in the present. Hence, any particle choosing in the present a possibility $a \in A$ must fortunately not go through all internal possibilities $b \in B$ simultaneously, as it is usually assumed in quantum theory.

Similarly, the possibility aBC means that there is some interaction with A in a in the present, but the interactions with B and C are not further specified. Hence, we can identify aBC with a. The other possibilities in the formulas above are interpreted in the same way.

Now, we have defined non-elementary possibilities in terms of subsets of the possibility space. But what are outcomes? Let us consider three examples. The first one is the classical experiment where we throw a die three times. This can be viewed as three identical machines $ABC = AAA$ in series, where each machine A is described by the set of possibilities $\{1, 2, 3, 4, 5, 6\}$, and the particle is identified with the person who throws the dice. There are no internal possibilities, and each elementary possibility, say abc, is an outcome and thus can become a fact. For example $abc = 666$ is the elementary possibility that all dies show 6. The possibility space coincides with the classical sample space

$$P = \Omega = \{abc : \text{ where } a \in A, b \in B, c \in C\}. \tag{19}$$

Hence, we have a classical experiment. For fair dice their probabilities are $1/6^3$.

Let us change this experiment such that the result of the second die, say B, is not detected. In other words, the results of B are internal. Then the outcomes are aBc, and thus differ from the elementary possibilities. Clearly, a change of the experimental set up changes the probabilities. For fair dice these probabilities are $1/6^2$.

For the double slit experiment, where no detectors are at the slits, both slits at the wall W describe internal elementary possibilities. In the present, a particle interacts with W in exactly one slit, which cannot be a fact, since it is not detected. Hence, only the subsets $sWd_m \in F_{SWD}$ define *outcomes*, and thus *facts* in the past.

More general, for the experimental set up $P = ABC$, when we assume internal possibilities $b \in B$, the sample space of outcomes is the set of subsets

$$\Omega = \{aBc: \; a \in A, c \in C\}\,. \tag{20}$$

All other subsets of P are not outcomes.

A *probability amplitude* is a mapping φ from the field of possibilities F_P into the set of complex numbers

$$F \to \varphi_F \in \mathbb{C} \tag{21}$$

that satisfies

$$|\varphi_F|^2 \leq 1, \; |\varphi_P|^2 = 1, \tag{22}$$

and for any countable set of pairwise disjoint possibilities $F_m \in F_P$ such that $F = \cup_m F_m$ it is

$$\varphi_F = \varphi_{\left(\bigcup_m F_m\right)} = \sum_m \varphi_{F_m}. \tag{23}$$

The latter is the rule for the *superposition of probability amplitudes*.

Two possibilities F and G are called *independent*, if both have no influence on each other. In accordance with Laplace experiments and classical probability theory, the probability amplitudes for independent possibilities are multiplied, that is,

$$\varphi_{F \cap G} = \varphi_F \, \varphi_G. \tag{24}$$

Thus, the *multiply-and-add rule* carries over to complex numbers yielding quantum mechanics. These rules allow to compute probability amplitudes for all outcomes of the sample space Ω, that is, for all classical elementary events.

With *Born's rule*

$$\textit{Prob(F)} = |\varphi_F|^2 \textit{ for all } F \in \Omega, \tag{25}$$

we obtain from the calculated probability amplitudes of the outcomes the classical probabilities. Then we can use Kolmogorov's rules for obtaining probabilities for non-elementary events, that is, for the subsets of the sample space Ω. Notice that we apply Born's rule only to complex amplitudes of outcomes, and not to a set of outcomes. The reason is that the square of a sum of magnitudes of complex numbers is not equal to the sum of squared magnitudes of complex numbers.

In summary, these rules serve to calculate the complex amplitudes for the outcomes allowing interference. Born's rule provides probabilities for all outcomes, and with Kolmogorov's rules we obtain classical probabilities for the non-elementary events. In most applications, the important and difficult task is the calculation of the probability amplitudes of the outcomes. The probability amplitudes of the outcomes are the relevant quantities, and sometimes it is difficult to calculate them.

Let us consider our experimental set up $P = ABC$. We assume that the possibilities of machine B are internal, and the possibilities $ab := abC$ and $bc := Abc$ are independent. Then the possibilities $ac := aBc$ are the outcomes. The value φ_{ab} is the probability amplitude that if the possibility $a \in A$ is chosen, then the possibility $b \in B$ will be chosen in the next step. The value φ_{abc} is the probability amplitude that firstly the possibility $a \in A$ is chosen, then the possibility $b \in B$, and finally $c \in C$. The other probability amplitudes are defined analogously. Since the elementary possibilities $\{abc\}$ are pairwise disjoint, formula (23) implies

$$\varphi_{ac} = \sum_{b \in B} \varphi_{abc}. \tag{26}$$

Since $ab \cap bc = abc$, from (24) we get Feynman's well-known formula (5)[34]

$$\varphi_{ac} = \sum_{b \in B} \varphi_{ab}\, \varphi_{bc}. \tag{27}$$

Since all possibilities of B are internal, the probability detecting a particle in a and c must take all routes abc into consideration, although the particle chooses only one route in the present. The *superposition of probability amplitudes* (26) and (27) is the sum of several complex amplitudes, one for each route. This allows cancellation of probability amplitudes, yielding the typical phenomena of *interference*.

The superposition of amplitudes for calculating the amplitude of an outcome occurs only if the experiment contains internal possibilities. If there are no internal possibilities, the outcomes coincide with the elementary possibilities abc, and for each outcome there is exactly one route. Cancellation of amplitudes, and thus interference, does not occur. This is the reason why we speak of *classical experiments*, if internal possibilities are absent.

The term possibility is a notion of the future describing possible interactions of a particle with machines in the present. Superposition means that the possibilities of machine A can be expressed in terms of the possibilities of B, and these can be expressed in terms of those of C. "Expressed" means that a particle, interacting in the present with machine A choosing the possibility a, chooses with probability amplitude φ_{ab} the possibility $b \in B$ afterwards, and so on. This is our unspectacular interpretation of the celebrated *superposition principle* as a property of future possibilities, not of facts or outcomes that happen at one moment. Notice the difference to the *superposition of probability amplitudes* (26), where we sum up complex numbers over internal possibilities.

Let us go through this probabilistic framework in terms of the double-slit experiment SWD. Firstly, the possibility space is

$$P = \{sad_m, sbd_m : \ s \in S, a, b \in W, d_m \in D\}. \tag{28}$$

The internal possibilities are the slits in W. Secondly, the sample space of outcomes has the form:

$$\Omega = \{sWd_m : \ s \in S, d_m \in D\}. \tag{29}$$

[34] See Section 2, Feynman [1948]

Thirdly, we use the multiply-and-add rule:

$$\varphi_{sad_m} = \varphi_{sa}\,\varphi_{ad_m}, \quad \varphi_{sbd_m} = \varphi_{sb}\,\varphi_{bd_m}. \tag{30}$$

These are disjoint possibilities, and the superposition of both amplitudes yields the amplitudes of the outcomes

$$\varphi_{sWd_m} = \varphi_{sad_m} + \varphi_{sbd_m}. \tag{31}$$

Inserting the concrete amplitudes for the possibilities, see also QUITE, we obtain the amplitudes for the outcomes. With Born's rule we obtain the probabilities of the outcomes, and using Kolmogorov's rules we can calculate the probabilities for the non-elementary events. This small recipe shows the unbelievable simplicity of explaining the double slit experiment within the framework of possibilities and the trinity of time. Later, we show some further aspects of slit experiments.

We started these notes with probability paradoxes and the unanswered question: What are probabilities, and how can we calculate probabilities? Laplace experiments and the principle of indifference gave us rules for calculating probabilities for some problems. But starting with Bertrand's paradoxes we have seen that these rules are by far not sufficient. In particular, interference phenomena cannot be described, since probabilities are non-negative numbers. **Now we have a very general formalism working with complex probability amplitudes assigned to possibilities that allow us to calculate classical probabilities of outcomes, including interference phenomena. This formalism is a key in our interpretation of quantum mechanics, the latter known as the most fundamental physical theory.** From our point of view, quantum theory is a timeless theory of the future that uses the geometrical properties of experimental setups for calculating classical probabilities of outcomes. Quantum theory and classical probability theory are not different probability theories that are in contrast, as sometimes mentioned. In our approach they complement one another. Quantum theory describes future interactions, probability theory describes momentary decisions or present interactions that happen at some moment, and classical mechanics describe the facts of the past, namely the specific interactions that have happened.

The generalization of these postulates to a large number of machines, say $A, B, C, D, \ldots.., K$, is straightforward and left as an exercise.

Keep in mind: The recipe for calculating probabilities:
Given an experimental setup:

1. Define the possibility space P, and the internal possibilities.
2. Define the sample space Ω of outcomes.
3. Calculate the probability amplitudes for the outcomes by using the *multiply-and-add rule*, that is, the probability amplitudes for disjoint possibilities are added (*superposition*), and the probability amplitudes for independent possibilities are multiplied.
4. Calculate the probabilities for the outcomes using Born's rule.
5. Calculate with Kolmogorov's rules the probabilities for the classical non-elementary events.

The possibility space P and the field of subsets F_P are defined similarly as in classical probability theory the sample space Ω and the related field of subsets of the sample space. Moreover, in quantum theory the *multiply-and-add rule* holds true for probability amplitudes as well. The essential difference is (i) that amplitudes are complex numbers, (ii) that possibilities and outcomes are different quantities, and (iii) that internal possibilities, responsible for interference, are essential. Quantum theory can be viewed as a calculus with complex numbers that delivers numerical probabilities for outcomes based on experimental setups. This calculus is not restricted to microscopic systems. In contrast, it is mainly based on macroscopic machines. Quantum theory and classical probability theory are not different probability theories, but complement one another. We speak of classical experiments, if internal possibilities are absent. **This recipe completes our formulation of probability theory and the fundamentals of quantum mechanics.** Feynman's path integral, one of the mathematical equivalent formulations of quantum mechanics, is an immediate consequence of this recipe. Experiments, classical or quantum ones, can be explained by using this recipe.

3.3 The Vector Representation

We have defined and discussed *outcomes* or *elementary events*, the *sample space*, *elementary possibilities*, and *internal elementary possibilities*[35]. But how can we represent outcomes and possibilities? Looking at the previous examples it follows that both, possibilities and outcomes, can be represented by numbers. For a dice or a coin toss the outcomes are natural numbers. Similarly, the outcomes of most other physical models can be counted and represented by natural numbers. In summary, any experimental setup can be characterized by a collection of numbers that have a specific meaning in the experimental context. These numbers form mutually exclusive alternatives. We call this the *number representation* of outcomes or possibilities. It is the most commonly used representation in physics.

Additionally, there are two other mathematically equivalent representations. Any number can be represented as a string of bits, called *register*. Commonly, a *bit* is defined as a quantity that can take only one of two values, 0 and 1. It is the basic unit in information theory. A frequent interpretation of a bit is a question at a system that has exactly two possible answers, say YES for 1 and NO for 0. The two values are interpreted as logical values TRUE or FALSE. Since we can represent any number by a register, it follows that any possibility of a finite possibility space can be represented by a finite series of bits, that is, of binary possibilities which are related to YES-NO questions. The same holds true for outcomes. Thus, in addition to any number representation we have an equivalent *register representation* of outcomes or possibilities. The register representation is the most commonly used representation in information theory.

Quantum mechanics, a theory of linear algebra, is usually described by using *Dirac's "bracket" notation*: each vector in a linear space is written in the form

$$|x\rangle \tag{32}$$

where x is a label for the vector, and the notation $|\cdot\rangle$ denotes a column vector called "ket". The conjugate transpose of this vector is written as

$$\langle x|, \tag{33}$$

and called "bra". It follows that the inner product of two vectors $|x\rangle$ and $|y\rangle$

$$\langle x|y\rangle = \sum_i x_i^* y_i \tag{34}$$

is represented as a "bracket". It consists of the bra part $\langle x|$ and the ket part $|y\rangle$.

It is very natural to think of a classical bit as a two-dimensional vector

$$|\psi\rangle = \psi_0|0\rangle + \psi_1|1\rangle, \quad \psi_0, \psi_1 \in \{0, 1\}, \quad \psi_0 + \psi_1 = 1, \tag{35}$$

[35] For more examples, details and applications see Jansson [2017]

such that the bit value 0 is represented in the form

$$|\psi\rangle = |0\rangle \quad \Leftrightarrow \quad |\psi\rangle = \begin{pmatrix}\psi_0\\ \psi_1\end{pmatrix} = \begin{pmatrix}1\\ 0\end{pmatrix}, \tag{36}$$

and the bit value 1 is written as

$$|\psi\rangle = |1\rangle \quad \Leftrightarrow \quad |\psi\rangle = \begin{pmatrix}\psi_0\\ \psi_1\end{pmatrix} = \begin{pmatrix}0\\ 1\end{pmatrix}. \tag{37}$$

This is the *vector representation* of a bit. The two mutually exclusive bit values are represented as orthonormal vectors.

A string of two bits leads to the registers 00, 01, 10, and 11. We can represent them as four-dimensional orthonormal vectors

$$\begin{aligned} &|\xi\rangle = \xi_{00}|00\rangle + \xi_{01}|01\rangle + \xi_{10}|10\rangle + \xi_{11}|11\rangle, \\ &\qquad \xi_{ij} \in \{0,1\}, \sum_{i,j=0}^{1} \xi_{ij} = 1. \end{aligned} \tag{38}$$

Thus, exactly one ξ_{ij} is equal to 1, the other ones are 0. The vectors corresponding to the four mutually exclusive register values are represented as the four-dimensional canonical orthonormal unit vectors in the complex vector space $\mathbb{C}^4$:

$$|00\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, |01\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, |10\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, |11\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{39}$$

If we write both bits in the form (35), that is,

$$\begin{gathered} |\psi\rangle = \psi_0|0\rangle + \psi_1|1\rangle, \quad |\varphi\rangle = \varphi_0|0\rangle + \varphi_1|1\rangle, \\ \psi_0, \psi_1, \varphi_0, \varphi_1 \in \{1,0\}, \quad \psi_0 + \psi_1 = 1, \quad \varphi_0 + \varphi_1 = 1, \end{gathered} \tag{40}$$

take their tensor product, and compare with (38), then we obtain

$$\xi_{00} = \psi_0\varphi_0,\ \xi_{01} = \psi_0\varphi_1,\ \xi_{10} = \psi_1\varphi_0,\ \xi_{11} = \psi_1\varphi_1. \tag{41}$$

Hence, the coefficients ξ_{ij} are just the products of the coefficients ψ_i and φ_j, and the four possibilities have the register and vector representations (39).

The *vector representation* of a register of n bits is, via the tensor product of the bits in the register, a standard unit vector in $\mathbb{C}^{2^n}$. It is almost obvious and easy to show that each finite-dimensional linear space can be embedded into a tensor product of two-dimensional spaces.

The transition from the number representation or the register representation to the vector representation is called *vectorization*. The vector representation is widely used in quantum theory, when working with bits and qubits.

Summarizing, we have three mathematical equivalent representations: the number, register, and vector representation. This has some important consequences, as shown in QUITE[36]. Among them are consistent definitions of states

[36] Jansson [2017]

and observables for various physical models, a vectorized classical mechanics, or a unified treatment of classical mechanics, classical probabilistic mechanics, and quantum theory via semimodules.

Experimental setups or machines are characterized by their mutually exclusive and empirically decidable *alternatives*, namely the possibilities, outcomes or elementary events. In the vector representation these alternatives form orthonormal unit vectors. Thus we call them *base states*. *States* are vectors defined as *superpositions*[37] of base states:

$$|\xi\rangle = \sum_i \xi_i |i\rangle, \qquad \xi_i \in \mathbb{C}, \quad \sum_i |\xi_i|^2 = 1, \tag{42}$$

where $\{|i\rangle\}$ denotes an orthonormal base, describing the mutually exclusive and empirically decidable alternatives of a machine in its vector representation. The orthonormality implies that the coefficients, namely the *probability amplitudes*, are inner products

$$\xi_j = \langle j|\xi\rangle. \tag{43}$$

In our approach, the possibilities of one machine can be expressed in terms of the possibilities of another machine, in the sense that a particle, interacting in the present with the first machine choosing possibility $|\xi\rangle$ interacts with the other machine by choosing possibility $|j\rangle$ with probability amplitude $\langle j|\xi\rangle$.

Born's rule states that the transition from a state $|\xi\rangle$ to some state $|j\rangle$ is the squared magnitude of the amplitude $\langle j|\xi\rangle$. There is not much choice to replace Born's rule. Given a normalized state $|\xi\rangle$ as defined in (42), the completeness of the orthonormal base $\{|i\rangle\}$ implies

$$\sum_i |i\rangle\langle i| = 1, \quad \sum_i |\langle i|\xi\rangle|^2 = \sum_i \langle \xi|i\rangle\langle i|\xi\rangle = \langle \xi|\xi\rangle = 1, \tag{44}$$

where the 1 in the first equation denotes the identity operator. The sum over all squared magnitudes $|\langle i|\xi\rangle|^2$ is one, and thus they can be interpreted as a probability. If we would choose any other expression for the probability, according to the Theorem of Gleason, these expressions would not add up to one.

Since states can be described as vectors, it follows that the change of states is described by matrices or linear operators. Therefore, quantum mechanics is a linear theory in contrast to classical mechanics. But it turns out that a vectorized classical mechanics is a linear theory as well, working with permutation matrices, a special class of unitary matrices[38]. Given wave functions and Schrödinger's equation in the number representation, then vectorization leads immediately to the language and machinery of matrix algebra with many applications. For example, vectorization yields Newton's equation of motion in matrix form[39].

---

[37] Formally, superposition means that with vectors, representing states, their linear combination is also a state.

[38] Jansson [2017]

[39] See also Sections 8 and 9, Feynman [1948]

3.4 Superposition in the Vector Representation

In this section, we write our probability theory in terms of the *vector representation*. When we use Dirac's "bra-ket" notation and express the possibilities as an orthonormal basis, we get back well-known quantum formulas.

The elementary possibilities of machines A, B, and C are mutually exclusive and form orthonormal bases $\{|a\rangle\}, \{|b\rangle\}$, and $\{|c\rangle\}$ of corresponding Hilbert spaces. The possibilities of machine A can be expressed in terms of the possibilities of B, and these can be expressed in terms of those of C, in the sense that a particle, interacting in the present with machine A choosing the possibility $|a\rangle$, chooses with probability amplitude φ_{ab} the possibility $|b\rangle$ afterward, and so on. This *superposition principle* can be written as

$$|a\rangle = \sum_{b \in B} \varphi_{ab}|b\rangle, \text{ where } \sum_{b \in B} |\varphi_{ab}|^2 = 1 \text{ and } \varphi_{ab} = \langle b|a\rangle. \tag{45}$$

The last equation follows from the orthonormality, and using *Dirac's "bra-ket" notation*, where the initial state is on the right hand side of the bracket and the final state is on the left hand side. We have represented possibility $\{|a\rangle\}$ as a vector in the Hilbert space generated by the orthonormal basis $\{|b\rangle\}$. In the same way, we can express each possibility $\{|b\rangle\}$ in terms of the basis $\{|c\rangle\}$:

$$|b\rangle = \sum_{c \in C} \varphi_{bc}|c\rangle, \text{ where } \sum_{c \in C} |\varphi_{bc}|^2 = 1 \text{ and } \varphi_{bc} = \langle c|b\rangle. \tag{46}$$

Therefore, $|a\rangle$ can be represented as a vector in the Hilbert space generated by the orthonormal basis $\{|c\rangle\}$, yielding

$$|a\rangle = \sum_{b \in B} \varphi_{ab} \sum_{c \in C} \varphi_{bc}|c\rangle = \sum_{c \in C} (\sum_{b \in B} \varphi_{ab}\varphi_{bc})|c\rangle. \tag{47}$$

Then equation (47) yields the probability amplitudes of the possibilities $|aBc\rangle$:

$$\langle c|a\rangle = \varphi_{ac} = \sum_{b \in B} \langle c|b\rangle \; \langle b|a\rangle. \tag{48}$$

which coincides with the *superposition of probability amplitudes* (27). If we assume that the possibilities provided by machine B are internal, then the outcomes are the possibilities $|ac\rangle = |aBc\rangle$, and for calculating the amplitude of an outcome we have to sum up over all amplitudes of the routes abc with $b \in B$.

The complete experiment can be expressed as the superposition of its outcomes:

$$|\xi_{ABC}\rangle = \sum_{a \in A} \sum_{c \in C} \varphi_{ac}|ac\rangle, \tag{49}$$

where $|\varphi_{ac}|^2$ is the probability of outcome $|ac\rangle$.

For the double-slit experiment, formulas (30) and (31) can be written as

$$\langle d_m|s\rangle = \langle d_m|a\rangle\ \langle a|s\rangle + \langle d_m|b\rangle\ \langle b|s\rangle. \tag{50}$$

Keep in mind: The *superposition principle* in the interpretation of the trinity of time is the unspectacular property of expressing possibilities of one machine in terms of the possibilities of other machines. Consequently, **a quantum state is not a property of one or more particles, but instead represents the properties of an experimental setup.**

These rules were never falsified when appropriately applied. A common ground in probability theory is the *multiply-and-add rule*: (i) in quantum theory the probability amplitudes for disjoint possibilities are added, and the probability amplitudes for independent possibilities are multiplied, (ii) in classical probability theory we use the multiply-and-add rule with nonnegative real numbers, and (iii) in Laplace experiments we count with natural numbers.

3.5 The Unbelievable Simplicity of Slit Experiments

In this section, although we already discussed some aspects of the double-slit experiment, we describe in more detail how simple and unspectacular slit experiments can be explained, when we use our probabilistic recipe. For some other aspects, see also Sections 2.6, 2.7, and 4.5 in QUITE.

The double-slit experiment[40] with its diffraction pattern has been called "The most beautiful experiment in physics". The used experimental setups depend on the type of objects interacting with the slit apparatus. It can be done with photons or electrons, and becomes more difficult for increasing size of the particles. Even large molecules, combined of 810 atoms, show interference. In 2012, scientists at the University of Vienna developed a double-slit experimental setup using large molecules called phthalocyanine. These molecules can be seen with a video camera. The molecules are sent one at a time through the wall with slits, such that exactly one molecule only interacts with the setup. At the screen of detectors they arrive localized at small places. This is typical for macroscopic objects, not for waves. Nobody has ever seen a collapsing wave. Moreover, the pictures produced with a video camera demonstrate the obvious incorrectness of the wave picture. But over a long period of time the molecules, one after the other, build up into an interference pattern consisting of stripes. This distribution shows the same kind of wave interference as if you drop two stones into a smooth pool, simultaneously. This seems to be evidence that this big molecule might travel as a wave. However, it is not a wave like a water or a sound wave. It is simply a probability distribution.

Strangely enough and frequently emphasized, the interference pattern also disappears if we obtain information only from one slit by using a detector there. The molecules passing through the other slit know what happens at the slit with detector. This phenomenon is called *non-locality*: what happens in one location seems to effect what happens in a distant location instantaneously. Non-locality is a fundamental mystery of today's quantum mechanics. Notice that in our approach this type of non-locality does not apply; the particle moves on one path in the present in agreement with the probabilities calculated with our recipe.

There is another strange mystery called the *observer effect*, that is, observing affects *reality*. Whether an interference pattern or a classical pattern occurs depends on observing the slits. The usual explanation is that "which-slit information“ makes the wave collapse into a particle. Therefore, in experiments we can change the way reality behaves by simply looking at it. Consequently, many physicists say that there is "no reality in the quantum world“. Below we show that the observer effect can easily be explained when carefully applying probability theory. There is no "wave collapse", but instead the detectors producing the "which-slit information“ change the whole experiment together with its sample space.

Moreover, there is another aspect of the widely celebrated *wave-particle duality*. In quantum mechanics, two-state systems are frequently discussed. These are systems that can exist in a superposition of two mutually exclusive

[40] Crease [2002]

base states. They form the fundamental quantities in quantum information theory, namely the *qubits*, or the *urs* as von Weizsäcker calls them. Polarization states or spin 1/2 states are examples. It is really questionable to use the term "wave" for a two state system. The right way is to speak of *probability distributions*, generated by all machines that form the experimental setup. These machines are globally positioned in a large area, and thus give the impression that quantum mechanics is non-local.

We mention a further series difficulty of the wave picture. Obviously, Schrödinger's wave equation can be no longer an ordinary wave propagating in spacetime, if systems with N particles are considered. Instead, it propagates in the so-called configuration space of dimension 3N, where even for a small macroscopic system this dimension becomes astronomically large.

Zeilinger, well-known for his pioneering experimental contributions to the foundations of quantum mechanics, gave an impressive talk in 2014 [41] "Breaking the Wall of Illusion". He said that in science we broke down many illusions in the course of history, for instance, that "the earth is flat", that "the earth is the center of the universe", that "we are biologically special and different from other animals", that "space and time is something absolute", and "in quantum mechanics we broke down many illusions about reality. One of the illusions, we first broke down in quantum mechanics, is that an object can only be at a given place at a given time. There have been many experiments about that. One of the experiments was done by Jürgen Blinek many years ago, the so-called double-slit experiment with atoms, which shows that atoms can go through two slits at the same time."

Can we resolve Zeilinger's quantum mysteries? My guess is yes. Our explanation of the double-slit experiment below breaks down the illusion that a particle is a wave and can be at several places at the same time. This supports experimental observations: an atom being at different places at the same time has been never measured. The latter statement is only a mathematical conclusion, not an experimental one. Secondly, the slit experiment (in our approach) is a simple consequence of a probability theory which carefully distinguishes between outcomes, possibilities, and internal possibilities. The experimental setup, consisting of various machines, is responsible for the patterns. These machines are distributed non-local over space. They are responsible for the possibility space, the sample space, and the probability amplitudes, thus for the concepts characterizing the future. The molecule's property are local interactions with the machines in the present. The patterns of the double-slit experiment are facts in the past. There is no mystery. Mystery occurs because most arguments are based on pushing the "local" properties of molecules in the foreground, and not the global aspects of experimental setups.

[41] see for example YOUTUBE

Keep in mind: The slit experiment in 2012 with the large phthalocyanine molecules shows: (i) a molecule is not a wave, (ii) supports clearly our probabilistic approach, (iii) the pictures of the molecules with the video camera show that a material object is not at different places at the same time, and (iv) it leaves many quantum interpretations at least doubtful.

We consider a source of objects. For the sake of convenience, we assume an experimental setup with a discrete space consisting of points

$$(m,t), \quad m = -2,-1,0,1,2, \quad t = t_0, t_1, t_2, \tag{51}$$

where m denotes five spatial points at three times t, as displayed in Figure 2. Of course, the future is timeless, and thus t represents no time. This notation means that the future events at t_0 would happen before the future events at t_1, and those before the future events at t_2. In other words, the values t_i describe merely a causal ordering of the experiment. It is a property of the ordering of machines inside the experimental set up, not time. This slit-experiment consists of three machines, the source S containing only one possibility s, the wall W with two slits a and b, and the screen of detectors $d_m \in D$. At t_0 a particle leaves the source S, arrives at t_1 at the wall with two slits a and b, and is detected at t_2 in exactly one of the detectors d_m. We ask for the probability that the particle is detected at point m.

It makes no difference for the mathematical treatment, if we choose a much finer grid, for instance with 10^{100} points, leading to an approximation of space-time that is much finer than the accuracy of any measurements.

Now we explain this experiment by using our recipe: firstly, we define the possibility space, secondly, we define the sample space, thirdly, the complex amplitudes are multiplied for independent possibilities and are added for mutually exclusive possibilities, and finally, for the outcomes the magnitudes of the amplitudes are squared according to Born's rule. In the following, we use the vector representation.

At first, we consider the experiment where b is closed. Then the possibility space is

$$P = \{|sad_m\rangle : \quad m = -2,-1,0,1,2,\}\,. \tag{52}$$

There are no internal possibilities. Therefore, the outcomes coincide with the possibilities, and the possibility space P is equal to the sample space Ω. Hence, according to our definition, we have a *classical experiment*. In the following we write shortly $|d_m\rangle$ for the outcomes.

The probability amplitude $\langle a|s\rangle = 1$, since we consider only particles that pass the slit a. Since the other slit is closed, it follows that

$$\langle d_m|s\rangle \;=\; \langle d_m|a\rangle\langle a|s\rangle \;=\; \psi_m. \tag{53}$$

There is only one route. Alternatively, we could set $\langle b|s\rangle = 0$. Then

$$\begin{aligned} \langle d_m|s\rangle &= \langle d_m|a\rangle\langle a|s\rangle + \langle d_m|b\rangle\langle b|s\rangle \\ &= \langle d_m|a\rangle \cdot 1 + \langle d_m|b\rangle \cdot 0 \\ &= \psi_m. \end{aligned} \tag{54}$$

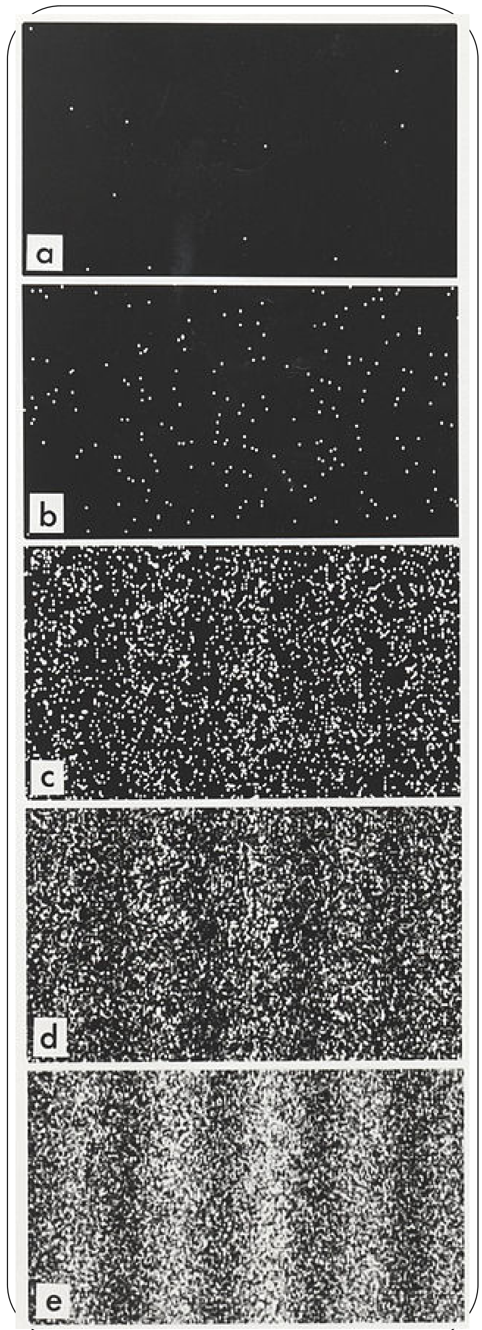


Figure 4: Results of a double-slit-experiment performed by Dr. Tonomura showing interference. The numbers of electrons are 11 (a), 200 (b), 6000 (c), 40000 (d), and 140000 (e). The electrons were shot one by one through the double-slit so that they could not interfere with each other.

Born's rule implies $\text{Prob}\{\langle d_m|s\rangle\} = |\psi_m|^2$. Thus, we obtain a classical probability without any interference, as expected.

Secondly, we consider the experiment where slit a is closed. Using the same arguments as above, we obtain $\text{Prob}\{\langle d_m|s\rangle\} = |\varphi_m|^2$ without any interference.

Finally, we assume that both slits are open. Then the possibility space is

$$P = \{|sad_m\rangle, |sbd_m\rangle : \quad m = -2, -1, 0, 1, 2, \}\,. \tag{55}$$

The internal possibilities are the two slits a and b in the wall W. Hence, the sample space consists of the outcomes

$$\Omega = \{|sWd_m\rangle : \quad m = -2, -1, 0, 1, 2, \}\,, \tag{56}$$

In the following, we write shortly $|d_m\rangle$ for these outcomes. The possibility space is larger than the sample space yielding a non-classical model.

We assume that the experiment is symmetric with respect to both slits, that is, in a future experiment the particles pass with probability $\frac{1}{2}$ through each slit. Hence, we set $\langle a|s\rangle = \langle b|s\rangle = \frac{1}{\sqrt{2}}$. The probability amplitudes calculated by the multiply-and-add rule are

$$\begin{aligned} \langle d_m|s\rangle &= \langle d_m|a\rangle\langle a|s\rangle + \langle d_m|b\rangle\langle b|s\rangle \\ &= \tfrac{1}{\sqrt{2}}\psi_m + \tfrac{1}{\sqrt{2}}\varphi_m. \end{aligned} \tag{57}$$

Therefore, we get the probabilities

$$\begin{aligned} \text{Prob}\{\langle d_m|s\rangle\} &= |\tfrac{1}{\sqrt{2}}\psi_m + \tfrac{1}{\sqrt{2}}\varphi_m|^2 \\ &= \tfrac{1}{2}(\psi_m + \varphi_m)^*(\psi_m + \varphi_m) \\ &= \tfrac{1}{2}(\psi_m^*\psi_m + \psi_m^*\varphi_m + \varphi_m^*\psi_m + \varphi_m^*\varphi_m) \\ &= \tfrac{1}{2}\left(|\psi_m|^2 + |\varphi_m|^2\right) + \tfrac{1}{2}(\psi_m^*\varphi_m + \varphi_m^*\psi_m). \end{aligned} \tag{58}$$

It follows that the first term in this sum corresponds to the classical probability, and the second term describes interference.

This can easily be seen as follows. For points m with $\psi_m = \varphi_m$ we obtain from (58)

$$\text{Prob}\{\langle d_m|s\rangle\} = 2|\psi_m|^2. \tag{59}$$

This doubles the classical probability, where only one slit is open. Hence, we have *constructive interference*. If $\psi_m = -\varphi_m$, the probability of finding the particle at point m is

$$\text{Prob}\{\langle d_m|s\rangle\} = 0, \tag{60}$$

yielding *destructive interference*. For other combinations we obtain probabilities that are between both extreme cases, as displayed in Figure 5.

Until now we don't have the correct values for all amplitudes, such as ψ_m and φ_m. In order to calculate the amplitudes for particle with momentum p going from one position x_1 to another x_2 we need the classical physical action

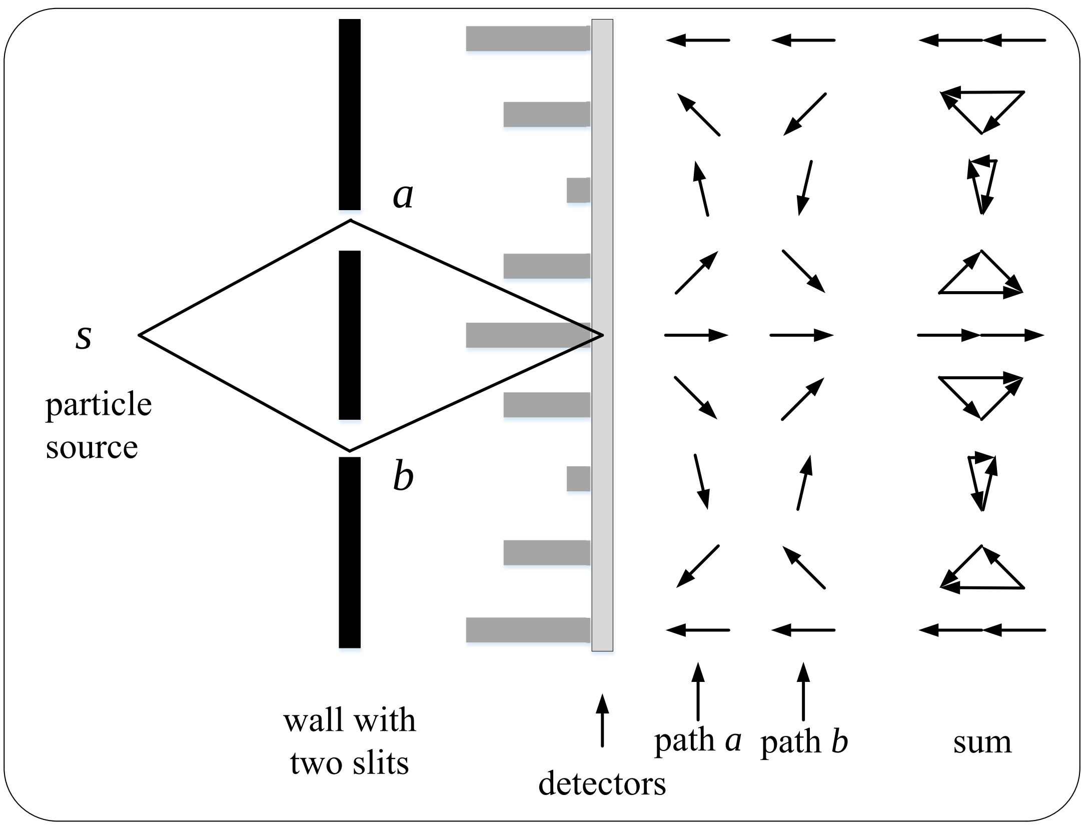


Figure 5: Schematic illustration of the double-slit experiment. The arrows represent the complex amplitudes for each path and their sum. Squaring the magnitude of the sum determines the corresponding probability. This leads to destructive and constructive interference, as displayed on the wall of detectors.

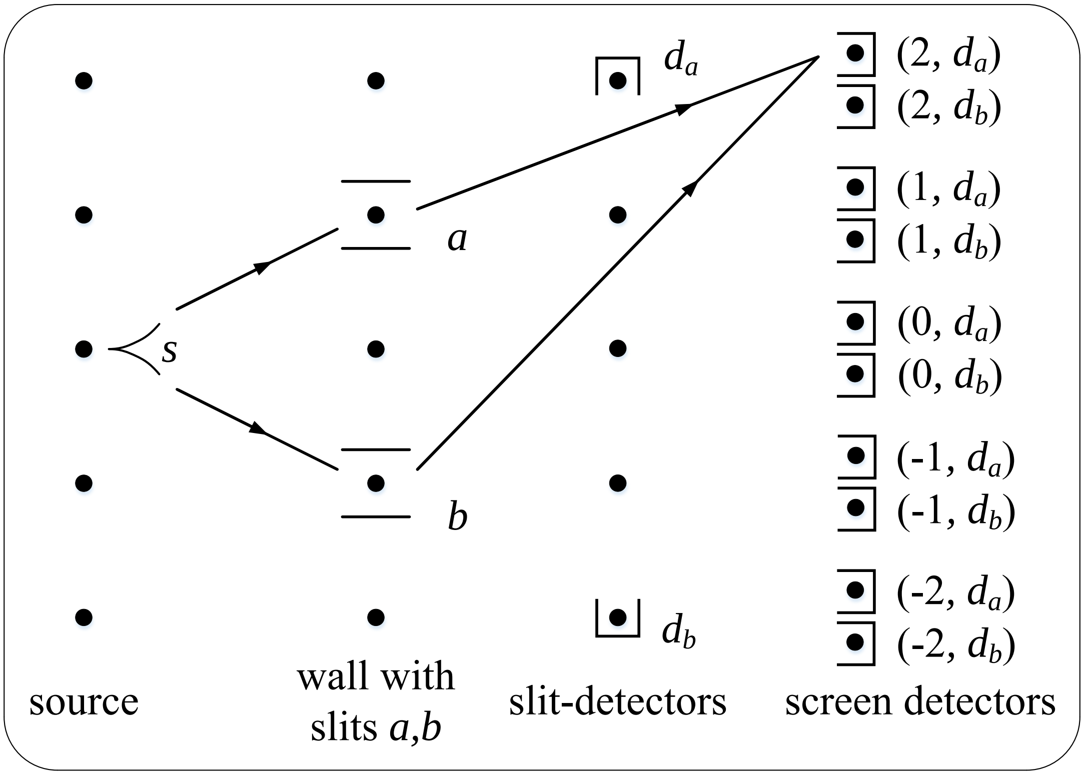


Figure 6: The double-slit experiment with slit-detectors. There are two paths for the event that a particle arrives at point 2 and detector d_a clicks. For the other points there are two paths as well.

of this process, usually denoted by S. In classical mechanics a first order approximation of the action is $S = p(x_2 - x_1)$, and the related amplitude of a path between positions x_1 and x_2 is proportional to the complex number

$$\langle x_2|x_1\rangle = e^{iS/\hbar} = e^{ip(x_2-x_1)/\hbar}, \tag{61}$$

where $\hbar$ is Planck's constant.

Now, we want to discuss the case where we can get information about which slit the particle passes through. This information can be given by two additional detectors d_a and d_b that click when a particle passes slit a or b, respectively. Of course, detectors may fail and information might be wrong. Such cases are not considered at the moment. We assume that the detectors work correctly, that is, it cannot happen that a particle arrives at m via slit b and detector d_a clicks, or both detectors d_a or d_b don't click.

The experimental setup has changed. Now, we apply our recipe. Additionally, we have at the third place the machine $I = \{d_a, d_b\}$ of detectors which gives information through which slit a particle passes. Looking at the experiment $SWID$, displayed in Figure 6, we have the possibilities that a particle is detected at point m and the detector d_a or d_b clicks. Obviously, there are no internal possibilities. Therefore, the outcomes coincide with the possibilities, and the possibility space is equal to the sample space:

$$P = \Omega = \{|sad_ad_m\rangle, |sbd_bd_m\rangle : \quad m = -2, -1, 0, 1, 2, \}. \tag{62}$$

Thus, we have a classical experiment without any interference. But the outcomes have changed. They are doubled. Of course, a change of the possibility space and the sample space must imply a change of the statistics.

In the following we denote $|m, d_a\rangle$ and $|m, d_b\rangle$ as the outcomes. The amplitude that a particle goes from source s via slit a to point m and detector d_a clicks is

$$\langle (m, d_a)|s\rangle = \langle m|d_a\rangle\langle d_a|a\rangle\langle a|s\rangle. \tag{63}$$

For each outcome we have exactly one path. Our assumptions imply the probability amplitudes $\langle d_a|a\rangle = 1$ and $\langle a|s\rangle = \frac{1}{\sqrt{2}}$. Hence,

$$\langle (m, d_a)|s\rangle = \langle m|a\rangle\langle a|s\rangle = \frac{1}{\sqrt{2}}\psi_m, \tag{64}$$

which leads to the classical probability $\text{Prob}\{\langle (m, d_a)|s\rangle\} = 1/2|\psi_m|^2$, see (54). Analogously, we obtain the classical probability $\text{Prob}\{\langle (m, d_b)|s\rangle\} = 1/2|\varphi_m|^2$. The fact that in this experiment internal possibilities are absent, such that possibilities coincide with outcomes, implies the disappearance of interference. It is simply a consequence that there are no internal possibilities. This has nothing to do with human observes, as frequently and erroneously mentioned.

The same result is obtained when we use only one detector, say detector d_a. Then the detector d_b is replaced by the possibility "detector d_a does not click". As above we obtain the same possibilities and outcomes yielding the same classical probabilities.

Finally, we consider the experiment where it may also happen that a particle arrives at m via slit b and detector d_a clicks, or that a particle arrives at m via slit a and detector d_b clicks. The case where both detectors click or both don't click is left as an exercise.

Now the slits represent internal possibilities. The possibility space is

$$P = \{|sad_ad_m\rangle, |sbd_bd_m\rangle|sad_bd_m\rangle, |sbd_ad_m\rangle : \quad m = -2, -1, 0, 1, 2, \}. \tag{65}$$

yielding the sample space:

$$\Omega = \{|sWd_ad_m\rangle, |sWd_bd_m\rangle : \quad m = -2, -1, 0, 1, 2, \}. \tag{66}$$

As before we denote $|m, d_a\rangle$ and $|m, d_b\rangle$ as the outcomes.

In our new experiment, the amplitude that a particle goes from source s via slit a to point m and detector d_a clicks is

$$\langle m|d_a\rangle\langle d_a|a\rangle\langle a|s\rangle. \tag{67}$$

But it may also happen that a particle arrives at m via slit b and detector d_a clicks. This possibility has the amplitude

$$\langle m|d_a\rangle\langle d_a|b\rangle\langle b|s\rangle, \tag{68}$$

and should happen rarely, provided the detectors work well. Other mutually exclusive possibilities do not occur, as can be seen from Figure 6. According to our addition rule we have to add both amplitudes

$$\langle (m, d_a)|s\rangle = \langle m|d_a\rangle\langle d_a|a\rangle\langle a|s\rangle + \langle m|d_a\rangle\langle d_a|b\rangle\langle b|s\rangle \tag{69}$$

for the outcome that a particle arrives at m from source s and detector d_a clicks. Thus, we have a non-classical model with interference. The corresponding probability is

$$\text{Prob}\{\langle (m, d_a)|s\rangle\} = |\langle (m, d_a)|s\rangle|^2. \tag{70}$$

If the detectors are perfect, then the probabilities

$$\text{Prob}\{\langle d_a|a\rangle\} = 1 \quad \text{and} \quad \text{Prob}\{\langle d_a|b\rangle\} = 0, \tag{71}$$

yielding $\langle d_a|a\rangle = 1$ and $\langle d_a|b\rangle = 0$. The superposition vanishes, and assuming $\langle a|s\rangle = \frac{1}{\sqrt{2}}$ gives

$$\text{Prob}\{\langle (m, d_a)|s\rangle\} = |\langle m|a\rangle\langle a|s\rangle|^2 = \frac{1}{2}|\psi_m|^2, \tag{72}$$

which is the classical probability as in (64).

With the same arguments as before, we obtain the amplitude

$$\langle (m, d_b)|s\rangle = \langle m|d_b\rangle\langle d_b|a\rangle\langle a|s\rangle + \langle m|d_b\rangle\langle d_b|b\rangle\langle b|s\rangle \tag{73}$$

for the event that a particle arrives at point m from source s and detector d_b clicks. For an ideal detector $\langle d_b|a\rangle = 0$, and assuming $\langle b|s\rangle = \frac{1}{\sqrt{2}}$, we obtain the classical probability

$$\text{Prob}\{\langle (m, d_b)|s\rangle\} = |\langle m|b\rangle\langle b|s\rangle|^2 = \frac{1}{2}|\varphi_m|^2. \tag{74}$$

Finally, if both detectors don't work, say all amplitudes $\langle d_a|a\rangle$, $\langle d_a|b\rangle$, $\langle d_b|a\rangle$, and $\langle d_b|b\rangle$ are equal to some value α, then the total amplitude becomes

$$\begin{aligned} \langle m|s\rangle &= \langle (m, d_a)|s\rangle + \langle (m, d_b)|s\rangle \\ &= \alpha \cdot (\tfrac{1}{\sqrt{2}}\psi_m + \tfrac{1}{\sqrt{2}}\varphi_m). \end{aligned} \tag{75}$$

Comparing with (57), it follows that this is the probability if both slits are open, except for the factor α. Thus we have interference as in the case without any detectors.

All of our calculations have nothing to do with properties of particles or human beings that observe. Particles must only interact with the experimental setup. Only the experimental setup, consisting of several machines, determines the possibilities, outcomes, and amplitudes. A change of outcomes and possibilities in the experiment implies a change of the probabilities. That's all, nothing strange happened.

In many textbooks, phrases like "a conscious observing implies that interference vanishes" can be found[42] . This popular belief, that a conscious mind can directly affect reality, is not necessary when using the trinity of time. We have a clear cut between interacting objects and experimental setups, distinguishing between possibilities, internal possibilities, and outcomes or elementary events. In our approach, last looker or other strange properties are not required.

Keep in mind: When using the concept "trinity of time" with distinguishing between possibilities, internal possibilities, and outcomes, then last looker or other strange properties are not required.

[42]The term "observed" is ambiguous. What means "observing a quantum system"? Is it an interaction with another physical system of matter or energy, say the measuring apparatus? Or does it require a human intelligent mind? In our approach, any *observations* or measurements of a quantum system are physical interactions with a measuring apparatus or the environment in the present.

3.6 Hardy's Paradox

Hardy's paradox[43], published in 1992, is a spectacular experimental setup, where simple logical arguments about its physical constraints lead to a surprising contradiction: logic says that the experiment is not realizable, although it is realized with photons. See also Laloë for a nice presentation[44] on this. Hence, this paradox is a challenge for each interpretation, also for our probabilistic approach.

Hardy's paradox involves a two-qubit state consisting of two particles that are emitted simultaneously from a source S. Each particle interacts with two machines in series, say A and A' for the first particle and B and B' for the second particle. They can choose two kinds of possibilities for each machine which we denote by the values ± 1, that is, $a = \pm 1$ for $a \in A$ and so on.

Let us, for instance, consider a polarization experiment. Imagine that the not-primed possibilities a and b describe interactions of two photons with birefringent plates such that the z-axis is the optical axis. The primed possibilities a' and b' are the possibilities with respect to a plate with an optical axis in the xz-plane making an angle ϑ with the z-axis. Horizontal polarization is denoted by the value -1, and vertical polarization is denoted by the value $+1$.

Let us now look at an optical experimental setup, where additionally the following three conditions must be fulfilled when performing this experiment several times:

(i) The not-primed result, $a = +1$ and $b = +1$, sometimes occurs.

(ii) The mixed-primed results, $a = +1$ and $b' = +1$ as well as $a' = +1$ and $b = +1$, never occur.

(iii) The doubly-primed result, $a' = -1$ and $b' = -1$, never occurs.

The first condition guarantees that, when performing the experiment several times, sometimes the result $a = +1$ and $b = +1$ is obtained. But in these cases the second condition implies that $b' = -1$, since we have $a = +1$. Analogously, it must be $a' = -1$, since we have $b = +1$. Obviously, this violates the third condition. In other words, the three conditions cannot be fulfilled, simultaneously, leading to a striking contradiction.

A real surprise is the existence of a well-working experimental realization of Hardy's thought experiment with photons[45]. How can it be, although logic tells us that there cannot be a realization? The basic building blocks of this experiment is a pair of Mach-Zehnder interferometers that interact through a beam splitter.

Hardy's paradox is not restricted to photons. In principle, we can also take spin 1/2 particles. There is a subtle, not essential, difference between both types of particles. The two orthonormal base states for photons are horizontal and vertical polarization with respect to some axis. The two spin base states are called "up" and "down", with the difference that the angles for photons are

[43] Hardy [1992]
[44] Laloë [2001]
[45] Irvine, et al. [2005]

mathematically handled as being half the angles for spin 1/2 particles. Apart from that, the mathematical framework for photons and spin 1/2 particles is the same.

Moreover, we can also describe and discuss Hardy's paradox in terms of a three-slit experiment, and we will do this in the following. The experimental setup consists of a source that produces simultaneously two particles, which arrive afterwards at a wall with three slits, and which are finally detected on a wall of position detectors. The realization must satisfy the following conditions:

(a) The value $a = +1$ denotes the possibility that the first particle passes in the present the topmost slit. The value $a = -1$ denotes the possibility that the first particle passes the middle slit.

(b) The value $b = +1$ denotes the possibility that the second particle passes in the present the lowest slit. The value $b = -1$ denotes the possibility that the second particle passes the middle slit.

(c) It is impossible that both particles pass the middle slit simultaneously. Perhaps, the slit is too small, or they annihilate each other. Moreover, the first particle cannot pass the lowest slit, and the second particle cannot pass the topmost slit, since there is a partition in the middle of the wall. For the possibility where both particles don't pass the middle slit in the present, this slit is open, and we have the usual interference patterns for each particle as in the double slit experiment. It follows that the final wall of detectors is partitioned into areas of destructive and constructive interference.

(d) The primed value $a' = +1$ denotes the possibility that in the present the first particle would end in the area of destructive interference. This can be the case only if the second particle blocks the middle slit, and thus destroys the area of destructive interference. The primed value $a' = -1$ is the negation, thus denotes the possibility that the first particle does not end in the area of destructive interference. In the same way, the primed values $b' = +1$ and $b' = -1$ are defined.

Now we have a thought experiment of a two-qubit state system which satisfies the conditions above. Does it work, and if yes, why? The condition (i) is satisfied, since it is possible that the first particle may pass the topmost slit whereas the second particle may pass the lowest slit, that is, "$a = +1$ and $b = +1$" may sometimes occur. The second condition (ii), that "$a = +1$ and $b' = +1$ never occurs" is fulfilled, since the first particle passes in the present the topmost slit such that the middle slit is free. Hence, the second particle comes to two open slits, and thus cannot end in the area of destructive interference. Similarly, the second statement of condition (ii) is fulfilled. Finally, the third condition (iii) is satisfied, because both interference patterns are not destroyed, if both particles can simultaneously pass the middle slit. But because of (c) this is forbidden, such that $a' = -1$ and $b' = -1$ never occurs.

Because of the contradiction proved above it seems that this slit experiment cannot be realized. What is wrong? However, we show that the experiment can be realized theoretically without any contradiction, and we explain it with classical logic by using our probability recipe.

Let us apply our recipe. Our first step is to find all possibilities. The elementary possibilities are the subsets $\{saba'b'\}$ consisting of one element. In the vector representation, these are all states

$$|saba'b' : a = \pm 1, b = \pm 1, a' = \pm 1, b' = \pm 1\rangle. \tag{76}$$

But many of them vanish because of the conditions above. For instance, the possibility where $a = -1$ and $b = -1$ vanishes, since both particles cannot pass the middle slit simultaneously. Because of condition (iii) the values $a' = -1$ and $b' = -1$ are not allowed. We simplify our notation slightly and write for non-elementary possibilities:

$$|a' = +1, b' = +1\rangle \quad = \quad |SABa'b' : a' = +1, b' = +1\rangle, \tag{77}$$

$$|a = +1, b' = -1\rangle \quad = \quad |SaBA'b' : a = +1, b' = -1\rangle, \tag{78}$$

$$|a = -1\rangle \quad = \quad |SaBA'B' : a = -1\rangle, \tag{79}$$

and so on.

Since there are no detectors at the slits, the internal possibilities describe the different situations in which both particles may pass the wall of slits. Therefore, the internal possibilities are

$$|a = +1, b = +1\rangle,\ |a = +1, b = -1\rangle,\ |a = -1, b = +1\rangle. \tag{80}$$

The possibility $|a = -1, b = -1\rangle$ is forbidden or has probability zero because of condition (c). Both particles can choose in a future experiment the first as well as the second and as well as the third internal possibility. Remember, if the experiment is performed in the present, then either the first or the second or the third internal possibility is chosen. Summarizing, we have described all possibilities and all internal possibilities.

In the second step we determine the sample space of outcomes. From conditions (i), (ii), and (iii) it follows that, in the vector representation, all outcomes are given by

$$|a' = +1, b' = +1\rangle,\ |a' = +1, b' = -1\rangle,\ |a' = -1, b' = +1\rangle. \tag{81}$$

The possibility $|a' = -1, b' = -1\rangle$ is forbidden or has probability zero because of condition (iii).

Since the complete experiment can be expressed as the superposition of its outcomes (49), it follows that the state vector of outcomes takes the form

$$|\xi\rangle = \alpha|a' = +1, b' = -1\rangle + \beta|a' = -1, b' = +1\rangle + \gamma|a' = +1, b' = +1\rangle. \tag{82}$$

This superposition of outcomes is an entangled state vector, where Born's rule applied to the probability amplitudes α, β, and γ yields the probabilities for the outcomes. We calculate the amplitudes in a moment.

Now, our basic question, whether Hardy's experiment is realizable, is equivalent to the question whether the probability of the possibility $|a = +1, b = +1\rangle$ is greater than zero. Hence, we ask for the existence of a non-zero probability amplitude for this possibility. If the probability amplitude is zero, then the experiment is not realizable.

Since the possibility $|a = +1, b = +1\rangle$ is not an outcome, this is in some sense an inverse problem. Thus in step 3 of our recipe, at first we calculate the probability amplitudes α, β and γ of the superposition (82).

The primed possibilities as well as the possibilities without prime form orthonormal bases in $\mathbb{C}^2$. We use a reference frame such that both vectors $|a = \pm 1\rangle$ and both vectors $|b = \pm 1\rangle$ are the unit vectors along the z-axis and the x-axis. The primed orthonormal vectors $|a' = \pm 1\rangle$ and both vectors $|b' = \pm 1\rangle$ define the basis with respect to a different axis in the xz-plane. This axis makes an angle ϑ with the z-axis. Now the important point is that in our approach we can express possibilities of one machine in terms of the possibilities of another machine, that is, we use *superposition* in terms of the Ansatz:

$$
\begin{aligned}
|a = +1\rangle &= \cos\vartheta |a' = +1\rangle + \sin\vartheta |a' = -1\rangle, \\
|b = +1\rangle &= \cos\vartheta |b' = +1\rangle + \sin\vartheta |b' = -1\rangle.
\end{aligned} \tag{83}
$$

Since the tensor product is bilinear, we simply can write the possibility $|a = +1, b' = +1\rangle$ in terms of the primed quantities as

$$|a = +1, b' = +1\rangle = \cos\vartheta |a' = +1, b' = +1\rangle + \sin\vartheta |a' = -1, b' = +1\rangle. \tag{84}$$

According to Born's rule, the transition amplitude from one state to another (in the same Hilbert space) is described by the inner product of both states. In our case the interesting amplitude is the inner product $\langle\xi|\ a = +1, b' = +1\rangle$. Since both vectors in the inner product are primed quantities according to (82) and (84), this inner product is well-defined, Condition (ii) says that "$a = +1$ and $b' = +1$" never occurs, thus has probability zero. Hence, the related probability amplitude must be zero, that is,

$$0 = \langle\xi|\ a = +1, b' = +1\rangle \tag{85}$$

$$= \langle\xi|\ \cos\vartheta |a' = +1, b' = +1\rangle + \langle\xi\ |\sin\vartheta |a' = -1, b' = +1\rangle \tag{86}$$

$$. = \gamma\cos\vartheta + \beta\sin\vartheta. \tag{87}$$

Analogously, the exclusion condition $a' = +1$ and $b = +1$ is equivalent to

$$0 = \langle\xi|a' = +1, b = +1\rangle, \tag{88}$$

$$= \gamma\cos\vartheta + \alpha\sin\vartheta. \tag{89}$$

Both equations are equal to zero. Division with $\sin\vartheta$ yields

$$\alpha = \beta = -\gamma\cot\vartheta. \tag{90}$$

Since the state vector (82) is not normalized until now, we can multiply with $\sin\vartheta$ and get

$$|\xi\rangle = -\cos\vartheta\ (|a' = +1, b' = -1\rangle + |a' = -1, b' = +1\rangle) + \sin\vartheta\ |a' = +1, b' = +1\rangle. \tag{91}$$

We have obtained the state vector of the outcomes with its probability amplitudes. Now, we can ask for the probability for condition (i), that is, whether the internal possibility "$a = +1$ and $b = +1$" sometimes occurs. According to Born's rule we must calculate the bracket

$$\langle\xi|a = +1, b = +1\rangle, \tag{92}$$

square its magnitude, and divide this expression by the square of the norm of state ξ. With (83) and (91) it follows that

$$\begin{aligned} \langle\xi|a = +1, b = +1\rangle &= \langle\xi|(\cos\vartheta|a' = +1\rangle + \sin\vartheta|a' = -1\rangle) \\ &\otimes (\cos\vartheta|b' = +1\rangle + \sin\vartheta|b' = -1\rangle)\ \rangle \\ &= \langle\xi|\cos^2\vartheta|a' = +1, b' = +1\rangle \\ &+ \langle\xi|\cos\vartheta\sin\vartheta|a' = +1, b' = -1\rangle \\ &+ \langle\xi|\cos\vartheta\sin\vartheta|a' = -1, b' = +1\rangle \\ &+ \langle\xi|\sin^2\vartheta|a' = -1, b' = -1\rangle \\ &= -\cos^2\vartheta\sin\vartheta \end{aligned} \tag{93}$$

We square the magnitude of this expression, divide it by the square of the norm of state ξ, and obtain the probability for the possibility $a = +1$ and $b = +1$:

$$\text{Prob(a=+1 and b=+1)} = \frac{\cos^4\vartheta\sin^2\vartheta}{2\cos^2\vartheta + \sin^2\vartheta}. \tag{94}$$

The maximum of this probability is about 0.09.

Summarizing, our probability recipe based on a careful distinction between possibilities, internal possibilities, and outcomes proves that the three conditions (i), (ii), and (iii) can be realized, in contrast to the proof based on simple logical arguments. **Why?**

If you look into the literature, then either the realization of this experiment is not further discussed, or sometimes erroneous arguments are presented. In his interesting presentation Moti[46] writes:

> *A lot of confusion has been said about this experiment. People discussed specific properties of antiparticles and annihilation.*

[46] https://motls.blogspot.com/2011/01hardys-paradox-kills-all-realistic.htm

> *However, the surprising content of this thought experiment - and real experiment - has nothing whatsoever to do with antimatter. It is just another example of the difference between the incorrect classical logic and the correct quantum logic (including the rules of entanglement).* Moti 2011

However, going through the derivation above once more, we have used only classical logical arguments. Quantum logic is not necessary. The key difference to deterministic classical mechanics is the fact that quantum mechanics is a probability theory of the future: possibilities of one machine can be expressed in terms of the possibilities of another machine, according to the superposition principle which is described as a simple mathematical equation. Classical mechanics, as a theory of the past, works with facts, and thus superpositions like (83) are not allowed. In our derivation, it is the *superposition principle*, applied to internal possibilities, which is responsible for the realization of this spectacular experiment. Then interference occurs leading to the realization of Hardy's paradox.

But we can ask whether this derivation works also for classical probabilities. Hence, we use the Ansatz

$$|a = +1\rangle \quad = \quad p|a' = +1\rangle + q|a' = -1\rangle, \tag{95}$$

$$|b = +1\rangle \quad = \quad p|b' = +1\rangle + q|b' = -1\rangle, \tag{96}$$

where p and q are non-negative real numbers such that $p^2 + q^2 = 1$. Moreover, the values α, β and γ in (82) must be non-negative, such that the squared values sum up to one. Proceeding as above, and dividing with q yields

$$\alpha = \beta = -\gamma \frac{p}{q}. \tag{97}$$

Hence, negative values occur yielding a contradiction. It shows that this experiment is not realizable with the formalism of classical probability theory, and in particular with classical mechanics.

In our derivation we didn't use imaginary numbers, only real ones. Thus, using our rules with real numbers already explains the realization of this experiment. The reason is that the axes of the machines are in the xz-plane. It is possible to generalize Hardy's paradox to an experimental setup that requires axes in the real three-dimensional space. Then we need imaginary numbers, like in polarization and spin experiments.

The quantum superposition only applies if internal possibilities occur in the experiment. In other words, the distinction between possibilities and internal possibilities can be viewed as the basic reason for the realization of Hardy's experiment. Internal possibilities remain unknown and are not given outside. This is in contrast to many, perhaps most other interpretations of quantum mechanics. There, it is allowed that a superposition of microscopic particles carry over to macroscopic objects. Look, for example, at von Neuman's measurement, at Schrödinger's cat, or at Wigner's many mind interpretation.

3.7 The Frauchiger Renner Paradox

In the well-known journal "Quantamagazine" an article[47] about the Frauchiger Renner thought experiment starts with the statement " A thought experiment has shaken up the world of quantum foundations, forcing physicists to clarify how various quantum interpretations (such as many worlds and the Copenhagen interpretation) abandon seemingly sensible assumptions about reality".

The spectacular title of the paper of Frauchiger and Renner[48] is "Quantum theory cannot consistently describe the use of itself". In particular, a no-go theorem is presented stating that three very natural-sounding assumptions cannot all be valid. Table 4 in this paper presents violations of these assumptions for almost all well-known quantum interpretations. For instance, it is argued that no single-world interpretation can be self-consistent. A single-world interpretation is each theory stating that for measurements actually just one outcome occurs. In fact, our approach is a single-world interpretation.

This paper is closely related to Hardy's paradox together with "Wigner's friend" arguments. With our knowledge now, it may be a funny and fascinating exercise to find out whether this paper is true, thus falsifying our approach, or classifying it as a "crackpot paper" as pointed out in motls.blogspot.com.

[47] New Quantum Paradox Clarifies Where Our Views of Reality Go Wrong, Dec 4, 2018

[48] Frauchiger, Renner [2018]

3.8 Pinball, Polarization and Spin

In quantum mechanics *spin* is known as an intrinsic form of angular momentum carried by particles. Spin is required in order to describe several experiments, such as the Stern-Gerlach experiment. A simple visualization, already suggested 1925 by Uhlenbeck and Goudsmit, is a particle spinning around its own axis. It turned out that spin is quantized, that is, it can have only a finite number of discrete values with respect to a given axis. Electrons, or more general fermions, have spin 1/2, that is, they have two discrete values $s+ = +1/2$ and $s- = -1/2$ for each axis. In orthodox quantum theory, spin is described by a *spacetime-spin wave function*

$$\psi(x, s, t). \tag{98}$$

The degrees of freedom are the continuous position coordinates and the discrete spin values. Usually, this wave function is written as a column vector, where the spin values are related to the indices of the vector. For electrons we get

$$\psi = \begin{pmatrix} \psi_0(x,t) \\ \psi_1(x,t) \end{pmatrix} = \psi_0(x,t)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \psi_1(x,t)\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{99}$$

This representation is called a *spinor*. It allows to work appropriately with Pauli matrices.

But now we have a big problem. Complex spacetime-spin wave functions describe in our approach possibilities belonging to the future and representing the stochastics of experimental set ups, but not momentary interactions of the particles with the setup. The latter happen in the present in accordance with the related probability amplitudes. In other words, spin should be part of a random variable, in physics called wave function, which describes properties of the macroscopic mechanical experiments, not in the first place intrinsic properties of particles. A particle has mainly one property, namely that it can interact in the present with the experimental setup.

In the same way, an electron or a silver atom should have no spin. Instead the Stern-Gerlach platform should be responsible for the spin stochastics. Actually, Ohanian[49] argued that "spin may be regarded as an angular momentum generated by a circulating flow of energy in the wave field of the electron. [...] This provides an intuitively appealing picture and establishes that neither the spin nor the magnetic momentum are *internal* - they are not associated with the internal structure of the electron, but rather with the structure of its wave field". The wave field is generated by a machine, for instance the Stern-Gerlach apparatus.

All this may be similar to the pinball game. The pinball machine and the player describe by its mechanical structure the entire stochastics. The ball merely has to be just round with an appropriate size for interacting with the pinball machine. The game, namely all interactions, takes place in the present in accordance with the stochastics, the latter being part of the future.

Similarly, the polarization of a photon may not be associated with the internal structure of the photon, but the optical elements are responsible for

[49] Ohanian [1986]

the polarization stochastics. The same holds valid for other particles. This, of course, raises a completely different picture on particle physics, and questions many statements. This seems counterintuitive and is against the faith of almost all physicists. But this is what you have to pay, if you describe quantum theory as a probabilistic theory of future events. To some extent, this turns particle physics upside down. Von Weizsäcker[50], who often emphasized the splitting of time into past, present, and future, solved this dilemma by introducing a temporal logic. Instead we go the direct route, but preserve classical logic.

[50] Wei88, Wei92, Wei06

4 Feynman Revisited

In this section, we present some important aspects of Feynman's publications[51]. From his principles many other quantum mechanical rules emerge in a rather natural way. A remarkable feature of Feynman's formulation is on the one hand its simplicity, and on the other hand its universal applicability. They were never falsified when correctly applied. This is in contrast to many other physical frameworks.

Even strange claims survive until now. For example, Dyson writes:

> *Thirty-one years ago [1948], Dick Feynman told me about his "sum over histories" version of quantum mechanics. "The electron does anything it likes," he said. "It just goes in any direction at any speed, forward or backward in time, however it likes, and then you add up the amplitudes and it gives you the wave-function." I said to him, "You're crazy." But he wasn't.*[52]

We mention that today Feynman's approach is an actual fundamental tool in *string theory*. For instance, superstring scattering amplitudes in the RNS picture have been calculated only with Feynman sums over histories. There, the usual quantum operator calculus would be ugly, perhaps would be impossible.

In this section, we describe Feynman's approach in terms of our probability theory. In particular, we show that Feynman's path integral, one of the mathematical equivalent formulations of quantum mechanics, is an immediate consequence of our recipe. It would be a good exercise to read his publication in parallel and compare it with our presentation.

[51] Feynman [1948, 1985]

[52] https://en.wikiquote.org/wiki/Freeman_Dyson

4.1 The Probability Amplitude for a Space-Time Path

The previous ideas about probability and probability amplitudes can be used for describing the motion of a particle in spacetime[53].

Usually, a uniformly moving three-dimensional space of positions (X, v) is viewed as a reference frame. In QUITE, Sections 4.13 and 4.14, we have argued that such a three-dimensional space of positions can be viewed as a position machine, or as a "train" X moving at constant speed v. This machine is characterized by the mutually exclusive possibilities of positions $x \in X$. Now, imagine that our experimental setup consists of a finite number of position machines. In the present, a particle can interact with these machines at points, say with coordinates $x_0, x_1, x_2, ...$ successively. From a classical point of view, the possibilities x_i define a (non-continuous) path at successive times $t_{i+1} = t_i + \epsilon$.

Now, our central question is: what is the probability that a particle starts in $x_a \in X$ and ends in $x_b \in X$? We shall solve this problem with the recipe as described in Section 3.2.

At first, we ask: what is our possibility space? Obviously, the elementary possibilities of our experiment consist of all paths with position coordinates $x_a, x_1, x_2, ..., x_b$ starting in x_a and ending in x_b. They define the *possibility space* P. The complex probability amplitude of such a path is frequently written as a function of the coordinates:

$$\varphi(x_a, ..., x_i, x_{i+1}, ..., x_b) = \varphi_{x_a ... x_i x_{i+1} ... x_b}. \tag{100}$$

We use both notations.

Assuming that only the starting point x_a and the final point x_b are detected, and the other points of the path are internal possibilities, the outcomes can be defined in the form

$$x_a x_b = x_a XXX...X x_b. \tag{101}$$

These subsets of the possibility space P describe a particle starting in x_a and ending in x_b in a future experiment, where all other points x_i are internal possibilities. Hence, according to equation (23), the probability amplitude for an outcome is

$$\varphi_{x_a x_b} = const \sum_{all\ paths\ x_a \to x_b} \varphi(..., x_i, x_{i+1}, ...). \tag{102}$$

From Born's rule we obtain the probability

$$\text{Prob}(x_a x_b) = |\varphi_{x_a x_b}|^2. \tag{103}$$

for this outcome.

In summary, we have applied our recipe to space-time machines and have obtained probabilities for outcomes. However, the formulas (102) and (103) are not complete, since we don't know the explicit values of the probability

[53] See also Section 3, Feynman [1948]

amplitudes $\varphi(..., x_i, x_{i+1}, ...)$, as well as the normalization constant. This will be considered in the next section.

Finally, we mention the case of more degrees of freedom. Then the position coordinate x_i will represent a set of coordinates $(x_i^{(1)}, x_i^{(2)}, ..., x_i^{(k)})$ describing a configuration with k degrees of freedom. A path is a sequence of such configurations.

4.2 The Calculation of Probability Amplitudes for a Path

In the following we show how to compute numerically the important probability amplitudes for outcomes[54] yielding via Born's rule probabilities. Then we can apply Kolmogorov's rules to calculate probabilities for subsets of the sample space.

We remember Laplace's rules:

- (*Unity outcome)*: If there are several outcomes all contributing equally, and it is agreed that neither seems favored over the other, all outcomes should be equally likely assigned with the unit 1.

- (*Addition rule)* The probability of an event is summing up over all outcomes contained in this event, where each term in the sum is equal to 1, and then dividing by a normalizing constant, namely the number of all possible outcomes of the sample space. In other words, probability is the ratio of the favored elementary events to the total possible elementary events.

In some sense quantum mechanics generalizes these rules from unities to complex numbers, all equal in magnitude. We formulate the *generalized principle of indifference* or *unity outcome* as follows:

- (*Generalized unity outcome*) All paths contribute with complex numbers φ, equally in magnitude, that is,

$$\varphi(..., x_i, x_{i+1}, ...) = e^{\frac{i}{\hbar}S(...,x_i,x_{i+1},...)}. \tag{104}$$

  The phase S of their contribution is called the *action* of the path.

From this point of view, quantum mechanics turns out to be a generalized probability theory of Laplace experiments using complex numbers. In other words, it shows the unbelievable simplicity of quantum theory as a generalized Laplace calculus delivering numerical probabilities. We must only find an appropriate action for the experimental setup.

Obviously, the action, from which the classical equations of motion of a system can be derived, plays a fundamental role in physics. What is action? In classical mechanics, it is a functional which takes the trajectory or path of some system as its argument and maps it on a real number. Action has the dimensions of energy multiplied by time or momentum times length. More generally, the action S may be viewed as a universal number that summarizes specific geometrical properties of the experimental setup. For simple experiments, like throwing a die, the action can be defined as the zero function yielding $e^{\frac{i}{\hbar}S} = 1$. Then Feynman's rules lead to Laplace's rules in this special case. In other words, classical experiments, which can be solved by the rules of Laplace, can be solved with Feynman's formalism by defining the action equal to zero. The transition from elementary knowledge about probability theory,

[54] See Section 4, Feynman [1948]

as taught in school, to quantum mechanics is simple, not strange, and quantum mechanics turns out to be a natural extension of classical probabilistic rules.

On trajectories satisfying the Euler-Lagrange equations the action S, delivering the phase, is extremal. Hence, its first derivative with respect to small path-perturbations is zero. Therefore, the probability amplitudes for the paths in the neighborhood of the classical trajectory and those nearby almost have the same direction. Hence, they add up without cancellation yielding a high probability of the classical trajectory or orbit. The neighboring paths include smooth and non-smooth zig-zag ones. Non-extremal orbits only have low probabilities, because the amplitudes mostly cancel each other out[55].

The law (104) can be viewed as the main result in Feynman's theory. It links the formulation of classical physics in the form of an extremal principle, namely the action, with the Schrödinger equation in quantum physics. In our language, it connects the past with the future, while Born's rule connects the future with the present, and the collapse marries the present with the past.

Spacetime paths are explained in classical mechanics. There, it is well-known that the classical *action* of a path $x(t)$ requires evaluating the integral

$$S\left[x(t)\right] = \int L(x, \dot{x})dt, \tag{105}$$

where the Lagrangian $L(x, \dot{x})$ is a function of position and velocity. We assume that the path $(..., x_i, x_{i+1}, ...)$ between two points is connected by a straight line. Since the integral is a sum, we obtain

$$S\left[x(t)\right] = \sum_i S(x_{i+1}, x_i), \tag{106}$$

where the value $S(x_{i+1}, x_i)$ is the classical action for the path on the small line (x_{i+1}, x_i). This value is the one making the action extremal on this line.

The chronological order, frequently used in quantum physics and also used in Feynman's article, is to write the previous point x_i on the right hand side, and the final point x_{i+1} on the left hand side. In order to obtain the same formulas as in Feynman's article, we use his ordering.

Now we look at the problem of how to calculate the sum (102) over all paths. In the following we consider only the one-dimensional case, since it is easy to generalize the equations to the multi-dimensional case. For each path we set

$$dx_i = x_{i+1} - x_i,\ \epsilon = t_{i+1} - t_i,\ i = 0, ..., N. \tag{107}$$

Now we apply Riemann integration yielding via (106) the *Feynman path integral*

$$\varphi_{x_a x_b} = \lim_{\epsilon \to 0} \frac{1}{A} \int \int ... \int e^{\frac{i}{\hbar}\left[\sum_i S(x_{i+1}, x_i)\right]} ... \frac{dx_{i+1}}{A} \frac{dx_i}{A} ..., \tag{108}$$

[55] For a proof see for example Egli [2004]

where the normalization factor A is introduced such that condition (22) holds true for the probability amplitude $\varphi_{x_a x_b}$. This is similar as for Laplace experiments, where we need the normalization number N. Frequently, *Feynman's path integral* is denoted in the form

$$\varphi_{x_a x_b} = \int e^{\frac{i}{\hbar} S(x(t))} \, Dx(t). \tag{109}$$

Now we have a formula for the probability amplitude for general actions, which must be calculated explicitly. For a free particle the Lagrangian contains only kinetic energy, that is, $L = m\dot{x}^2/2$. A lot of calculations which can be found in most textbooks about quantum mechanics lead to

$$\varphi_{x_a x_b} = \left[\frac{2\pi i \hbar (t_b - t_a)}{m} \right]^{-1/2} \; exp \frac{im(x_b - x_a)^2}{2\hbar(t_b - t_a)}, \tag{110}$$

where $t_b - t_a$ denotes the required time. For other examples we refer the reader to the literature.

Finally, Feynman[56] writes that equation (108) completes his formulation of quantum mechanics. Therefore, all basic notions of quantum mechanics (wave functions, Schrödingers wave equation, operator algebra, Newton's equations, the Hamiltonian formalism, commutation relations, ...) can be mathematically derived from his path integral formalism. All these mathematical derivations can be found in his nicely written article.

We have derived Feynman's equation (108) by simply using our probabilistic recipe. Hence, this recipe can be viewed as the foundation of quantum theory. Moreover, we have shown that Feynman's formulation can be derived easily from a generalization of Laplace's rules and the *generalized principle of indifference*. This generalization is based on complex numbers, the largest field of reasonable numbers according to Hurwitz[57]. In this sense, quantum theory has nothing to do with strange concepts such as the wave-particle duality. It is a simple probability theory that applies everywhere, and is not restricted to microscopic systems, as mentioned several times.

> **Keep in mind:** Quantum mechanics turns out to be a generalized probability theory of Laplace experiments using complex numbers, the largest field of reasonable numbers according to Hurwitz. It shows the unbelievable simplicity of quantum theory, not strange, and easy to teach already in school.

[56] See end of Section 4, Feynman [1948]

[57] Jansson [2017, Section 2.2]

4.3 Schrödinger's Wave Equation

In this section we give a sketch of the derivation of Schrödinger's wave equation, thus proving that Feynman's formulation implies the ordinary formulation of quantum mechanics. All details can be found in his paper[58].

Since integration is summation, for any point x_c on a path between x_a and x_b the action integral satisfies

$$S(x_b, x_a) = S(x_b, x_c) + S(x_c, x_a). \tag{111}$$

Therefore, Feynman's path integral fulfills the equation

$$\varphi(x_b, t_b; x_a, t_a) = \int \varphi(x_b, t_b; x_c, t_c)\ \varphi(x_c, t_c; x_a, t_a)\ \frac{dx_c}{A}, \tag{112}$$

where $t_a < t_c < t_b$ denotes the times related to the positions x_a, x_c, x_b. Since the probability amplitudes for disjoint possibilities are added, and the probability amplitudes for independent possibilities are multiplied, the total amplitude $\Psi(x_b, t_b)$ to arrive at the point (x_b, t_b) is equal to the sum over all possible values x_a of the amplitudes $\varphi(x_b, t_b; x_a, t_a)$ multiplied by the total amplitude $\Psi(x_a, t_a)$. Thus, we obtain

$$\Psi(x_b, t_b) = \int \varphi(x_b, t_b; x_a, t_a)\ \Psi(x_a, t_a)\ \frac{dx_a}{A}. \tag{113}$$

This distribution of probability amplitudes is called *wave function*, although it has nothing to do with a wave. It's simply a probability distribution describing future events; nothing happens in contrast to classical waves.

We replace a and b by points x_k and x_{k+1}, respectively. Moreover, let $t = t_a$ and $t + \epsilon = t_b$ for small ϵ. Then we can write

$$\Psi(x_{k+1}, t + \epsilon) = \int \varphi(x_{k+1}, t + \epsilon; x_k, t)\ \Psi(x_k, t)\ \frac{dx_k}{A}. \tag{114}$$

Comparing with (108) we get Feynman's formula (18)[59]:

$$\Psi(x_{k+1}, t + \epsilon) = \int e^{\frac{i}{\hbar}[S(x_{k+1}, x_k)]}\ \Psi(x_k, t)\ \frac{dx_k}{A}. \tag{115}$$

Actually, our discrete path, and thus this equation, are only approximations of first order in ϵ. The following equations must be understood in this sense. We use the first order approximation of the action on the line (x_{k+1}, x_k):

$$S(x_{k+1}, x_k) = \epsilon\ L\left(x_{k+1}, \frac{x_{k+1} - x_k}{\epsilon}\right). \tag{116}$$

Let us consider a particle with kinetic energy $m\dot{x}^2/2$ and potential energy $V(x)$. The Lagrangian is the difference between kinetic and potential energy, thus yielding the approximation

$$S(x_{k+1}, x_k) = \frac{m\epsilon}{2}\ \left(\frac{x_{k+1} - x_k}{\epsilon}\right)^2 - \epsilon V(x_{k+1}). \tag{117}$$

---

[58] See Section 5 and the following ones, Feynman [1948]

[59] See p. 14, Feynman [1948]

Inserting this term into equation (115), some mathematical calculations[60] immediately lead to the well-known *Schrödinger equation*:

$$-\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} = \frac{1}{2m}\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2 \Psi + V(x)\Psi. \tag{118}$$

The orthodox *Copenhagen interpretation* says that a system of particles is completely described by its wave function Ψ, which lives in a Hilbert space. It obeys a linear first order partial differential equation, namely the Schrödinger equation. These are abstract mathematical statements without any relationship to sense experience, see also the various comments of well-known physicists about the quantum postulates in the introduction of QUITE. Frequently, in the Copenhagen interpretation, it is emphasized that quantum mechanics makes only statements about outcomes of experiments, but it is not allowed to ask what is actually happening. These statements have led to endless discussions.

Our probabilistic approach generates quantum mechanics from our experience of mutually exclusive alternatives which characterize not particles, but machines and the trinity of time, without leading to well-known paradoxes.

[60] See p. 20, Feynman [1948]

4.4 The Hamiltonian

Hamiltonian mechanics is a mathematical formalism that provides a deeper understanding of classical mechanics and of quantum mechanics[61]. In contrast to space-time with four independent variables, Hamiltonian mechanics works with six independent variables, namely three spatial coordinates x and three velocity coordinates v, or, if you like, three momenta coordinates p . In classical mechanics a *state* of a particle is a vector (x,p) of these six coordinates. The *phase space* is the set of all states. Paths in the phase space consist of sequences of states (x,p), that is, of pairs of positions and momenta or velocities. A path can be visualized as a particle that jumps from one "train" (X, v_i) to another one.

The *Hamiltonian function*

$$H(x,p) = K(p) + V(x) \tag{119}$$

is the sum of kinetic energy and potential energy. The *Hamilton equations* are defined on the phase space:

$$\dot{p}_k = -\frac{\partial H}{\partial x_k}, \quad \dot{x}_k = \frac{\partial H}{\partial p_k}, \tag{120}$$

The usual interpretation of these beautiful symmetric equations is: given any Hamilton function $H(x,p)$ and the values of the position and momenta coordinates at some time t, the equations (120) give the coordinates of the resulting path at an infinitesimal time later. The complete trajectory in the phase space is obtained by successively updating the coordinates. This is the usual interpretation of Hamilton's equations. But in our interpretation these equations describe facts of the past, and we know that the past is timeless. What does t mean? Well it's just a geometrical parameter that allows us to describe the solution of the equations in an explicit form.

The relationship between Hamiltonian mechanics and the Lagrangian formalism is given by the equation

$$L(x,v) = pv - H(x,p), \tag{121}$$

where for a particle of mass m the kinetic energy is $K = mv^2/2$ and the momentum is $p = mv$. Thus Feynman's path integral can be defined with the Hamilton function instead of the Lagrangian[62]. In QUITE, Sections 4.13 and 4.14, many advantages of describing classical physics and quantum physics in terms of the phase space are explained.

[61] See also Section 10, Feynman [1948]
[62] Soff [2002]

4.5 QED

In this section we present a very short introduction into a rather difficult topic, namely the theory of quantum electrodynamics (QED). This is the oldest and the most successful quantum field theory. The other theories in particle physics are modeled in similar ways.

Although the details of this theory are complicated, some basic principles can be explained. A fundamental and very recommendable book with title "QED, the strange theory of light and matter" is written by Feynman[63]. We recommend the reader to work through this book parallel, in particular through Chapter 3. This would be beneficial, because our language of future events differs from the conventional reasoning in physics. It turns out, however, that our recipe for probability theory is a good starting point for a better understanding of QED.

QED describes the interaction of light and electrons, and thus all of chemistry and biology. In particular, the behaviour of electrons in molecules, light reflection, the Pauli principle, and all other well-known quantum phenomena can be explained with this theory, except gravitation and nuclear phenomena. Actually, this theory and consequently Nature deals with probability amplitudes, and Feynman[64] points out:

> *There are no "wheels and gears" beneath this analysis of Nature;*
> *if you want to understand Her, this is what you have to take.*

It is out of our scope to describe complicated formulas and details. But the fundamental ideas can be presented on few pages. QED works with the electromagnetic field, thus necessarily requires the relativistic (3+1)-spacetime with coordinates (x, t). There is a difference between time t and space. In space a material object can go forward or backward, left or right, and up and down. But it cannot visit yesterday. In other words, it cannot go backward in time. Unfortunately, in our approach the future, and therefore quantum theory, is timeless. How is it possible to preserve QED, our best theory?

Well, we use spacetime coordinates only for calculating amplitudes, and we can view spacetime formally as a machine itself described by four-dimensional coordinates (x, t) such that particles can move free with respect to spatial coordinates x, but only forward with respect to the coordinate t. In other words, space and time are not separate quantities with distinct meanings. Instead, they represent just different directions. Symmetry requires that there is another type of particles which move only backwards with respect to coordinate t, but move free in space X. These particles are called *antiparticles*. In summary, the coordinate t has a purely geometric character suitable for a timeless theory.

In the following we consider only two examples. Firstly, we ask for the probability that an electron and a photon, starting at two different points in spacetime, are detected later on two other points. Secondly, we investigate how an electron is scattered by an electromagnetic field. In QED such a field

[63] Feynman [1985]
[64] Feynman [1985, page 78]

can be described as a machine composed of a collection of photons. Electrons can interact with this machine. Thus, we have the actors, namely electrons and photons, and we ask how to describe their interactions.

Surprisingly, it turns out that in QED only three key actions are sufficient to describe all phenomena of light and electrons, namely the interactions:

- $p(x, y)$: A photon goes in spacetime from position x, t_x to position y, t_y.
- $e(x, y)$: An electron goes in spacetime from position x, t_x to position y, t_y.
- $s(z)$: A photon is scattered in spacetime by an electron, that is, the electron emits or absorbs a photon at position z, t_z, the latter called a *vertex*.

Notice that these actions do not happen in the future, they describe only possibilities. It's the same as if you say that a coin has head and tails, but you don't throw the coin. It turns out that all possibilities, internal possibilities and outcomes can be defined with these actions. The antiparticles can perform the three key actions as well.

In order to get short formulas we suppress frequently the t value such that we write only x instead of (x, t_x) . In the following figures, the space coordinate x is displayed on the horizontal axis, the coordinate t on the vertical axis, a traveling photon is represented as a wiggly line, a traveling electron is displayed as a straight line, and a scattering is drawn as a vertex where a photon, an incoming electron, and an outgoing electron meet, see Figure 7. Our scaling is not in terms of seconds. Instead we use a 45 degree angle for motion with the speed of light. Then the horizontal distance $y - x$ is equal to the vertical distance $t_y - t_x$.

With these three actions we can define all possibilities, including the internal ones. Each possibility has an amplitude which can be calculated according to our rules of the recipe. Feynman's formulas are mathematically expressed in terms of the spacetime coordinates. Since it is much easier to calculate with the corresponding Fourier transforms leading to energy and momentum coordinates, almost no textbooks use Feynman's formulas. In our short presentation we abstain to derive and write down the exact formulas for these amplitudes[65]. For the interested reader we recommend the book of Griffiths[66] about particle physics; see also the article of Kummericki[67].

The formula for the amplitude of the first action $\phi_{p(x,y)}$ is simple. It depends on the differences $y - x$ and $t_y - t_x$ where the essential contribution occurs for the speed of light. **Surprisingly, in Feynman's theory about quantum electrodynamics[68] a photon can move faster or slower than the speed of light.** There are nonzero amplitudes for these cases. It is allowed that an electron or a photon can move to any place in spacetime with

[65]They depend on polarization and spin of the particles. Essentially, they are the solutions of the Dirac equation for electrons and the Klein-Gordon equation for photons.

[66]Griffiths [2004]

[67]Kumericki [2016]

[68]Feynman [1985, page 89]

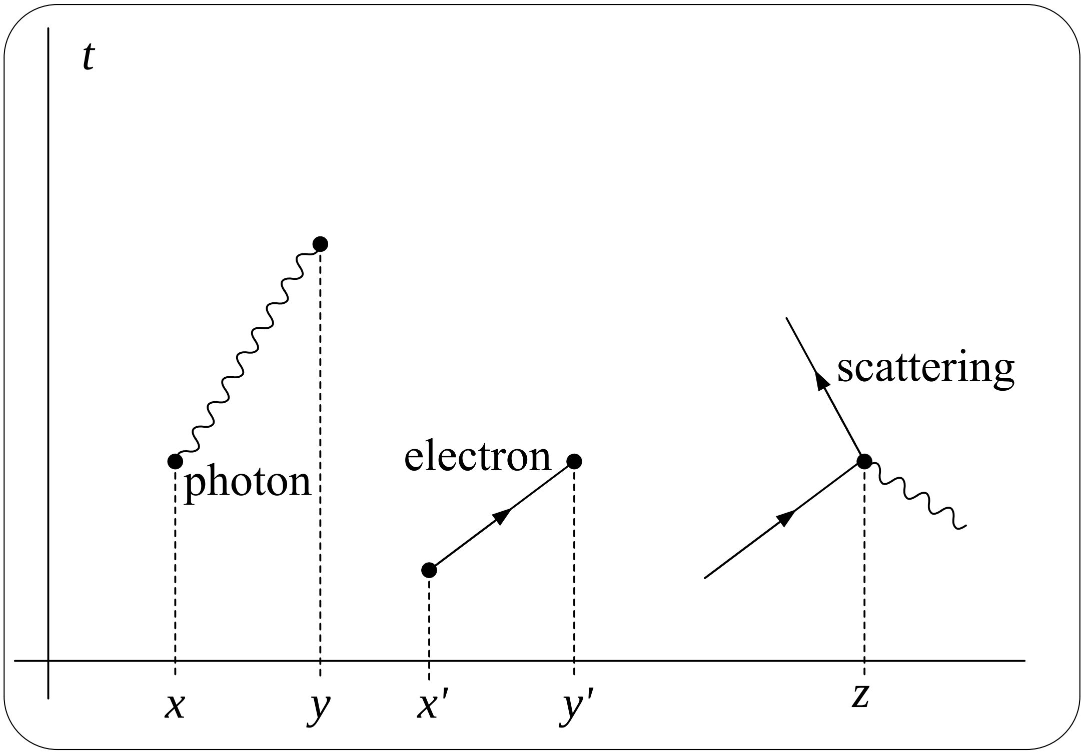


Figure 7: Feynman diagrams for the three key actions.

speeds greater than light. The mathematical formalism shows that these amplitudes almost cancel out over long distances. But over very short distances the photon travels zigzagging with speeds different to the speed of light.This contradicts the fundamental postulate in the theory of relativity, namely that nothing can move faster than the speed of light. Fortunately, **in QUITE, Chapter 4, we have derived the Lorentz transform, and thus the mathematical formalism of special relativity, without this postulate.**

Now we come to the second fundamental action $e(x, y)$. The corresponding amplitude $\phi_{e(x,y)}$ depends also on the differences $y - x$ and $t_y - t_x$, but is more complicated.

Finally, the third basic action $s(z)$ describing scattering is simple. Its amplitude $\phi_{s(z)}$ is just a number called j with value about -0.1.

With this preparation we can consider two examples, namely *Compton scattering* and calculating the value of the *magnetic moment*. We shall use our probabilistic recipe.

Compton scattering is a quantum process associated with an incoherent interaction between a single photon and a free electron. An electron and a photon are placed at distinct points x and y in spacetime, respectively. We ask for the probability of detecting the electron at u and the photon at v.

In the first step of our recipe we have to determine all possibilities including the internal ones. It suggests itself to consider all processes that can be constructed by using our three key actions.

The simplest type of possibilities is described by two paths between y and u for the electron and x and v for the photon, see the left diagram in Figure 8.

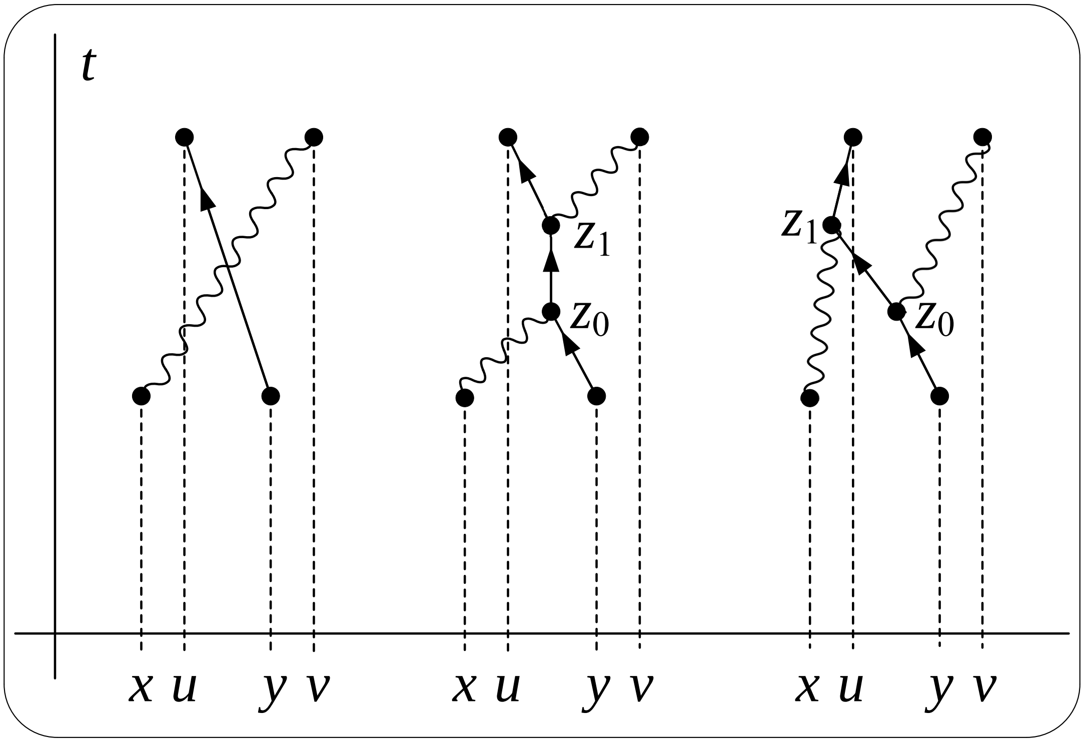


Figure 8: Compton scattering.

Both are independent , so that we have to multiply both amplitudes, obtaining $\phi_1(y,u)\cdot\phi_1(x,v)$. But that's not all. As we know from slit experiments, we have to compute all arrows for all possible paths between y and u for the electron and x and v for the photon. All these arrows represent mutually exclusive internal possibilities. Therefore, we have to add all these arrows yielding the amplitude

$$\phi_1 = \sum_{all\ paths\ y\to u, x\to v} \phi_1(y,u)\cdot\phi_1(x,v) \tag{122}$$

As we know, only the amplitudes of the paths that are infinitesimally close to the path of least action contribute constructively, whereas all arrows corresponding to the paths distant from the classical path of least action cancel out.

But the electromagnetic field, the machine consisting of a large, perhaps infinite, number of photons, allows everything what is not forbidden. So the next type of possibilities is displayed in the middle of Figure 8. The electron travels to vertex z_0, absorbs a photon, then travels to z_1 emitting a photon, and is finally detected at u, while the photon is detected in v. The vertices describe internal possibilities. The amplitudes are again calculated with the amplitudes of the three key actions: three electron actions, two photon actions and two scattering vertices. These independent actions must be multiplied for obtaining the amplitude describing two vertices. Then as before, we have to add up all arrows for all vertices z_0 and z_1 and all paths, obtaining the amplitude ϕ_2. This involves integration.

We have a third diagram on the right hand side of Figure 8. The electron travels to the internal vertex z_0, emits a photon detected in v, while the electron travels to the internal vertex z_1, where it absorbs the first photon, and is finally detected at u. Again the amplitudes are calculated as before, obtaining ϕ_3.

But this is not all. There is an infinite number of other possibilities with more and more internal vertices, including also antiparticles. They can be described by corresponding Feynman diagrams. Yes, the electromagnetic field is a magical machine making almost everything possible.

All amplitudes must be added. Diagrams with more internal vertices are multiplied with higher power of $j \approx 0.01$. This is a very small number, such that they don't contribute to the overall amplitude ϕ for the outcome, namely that an electron and a photon are placed at distinct points x and y in spacetime, and are detected at u and at v. Thus, the overall amplitude is

$$\phi = \phi_1 + \phi_2 + \phi_3 + ... \tag{123}$$

Squaring the magnitude of this amplitude yields the probability of the outcome. **This is the fundamental approach of QED. It can be derived with our probabilistic recipe in a natural way close to the Laplace concept.**

There is, however, a drop of bitterness. If we want to calculate amplitudes for couplings and consider all possible vertices that can occur until zero distances between them, then we obtain meaningless answers, since infinities occur. Actually, physicists obtained infinities for every problem in QED. Now, there exists a collection of techniques in QED and in quantum field theories, called *renormalization*, that are used to treat such infinities. For an introduction we recommend the books of Feynman and Griffiths[69].

Now we come to the calculation of the value of the magnetic moment. This number stands for the response of an electron to a magnetic field. More precisely, we want to compute the probability that an electron travels from one position to another and scatters with a photon. The positions of the electron and the photon are detected, thus describe the outcomes. A diagram is displayed in Figure 9.

It was Dirac who calculated the magnetic momentum by deriving a formula for the amplitude $\phi^1_{e(x,y)}$. He obtained the value 1 when using appropriate units. But later experiments showed that this number was not right. Instead it should be about 1.00116.

Well, assuming that this measured amplitude is a better approximation, the magical field of photons suggests the existence of other possibilities. Another possibility is a path where the electron emits a photon at vertex z_1, then scatters with a photon at z_0, and finally absorbs a photon at spacetime point z_2 before reaching the final position u, see Figure 10.

Therefore, electron lines and photon lines meet at three vertices z_0, z_1, and z_2 yielding an amplitude proportional to the small value j^3. More precisely, we write

$$\phi^2 = j^3 \int \psi_{e(1,0)}\psi_{e(0,2)}\psi_{p(1,2)}dz, \tag{124}$$

[69]Griffiths [2004], Feynman [1985]

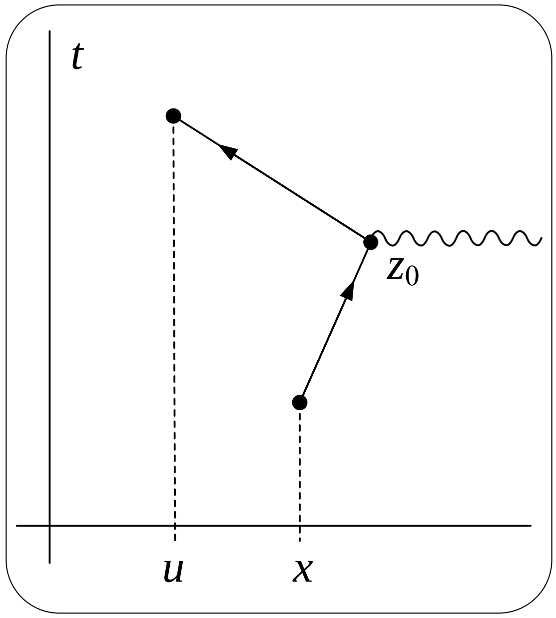


Figure 9: Feynman's diagram for calculating the magnetic moment.

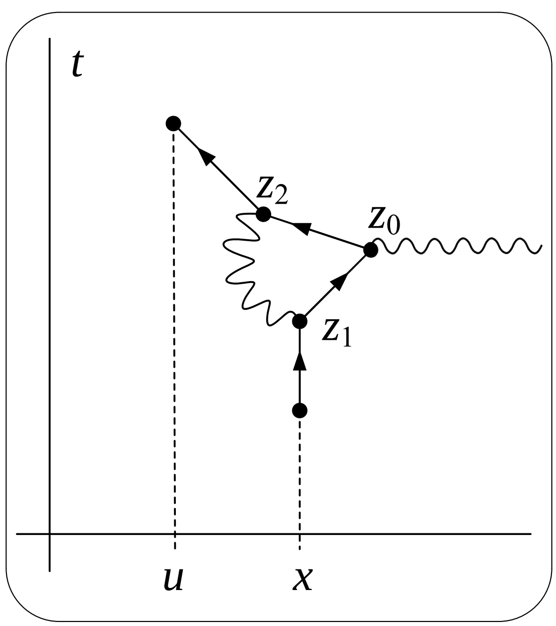


Figure 10: Another possibility for improving the magnetic moment.

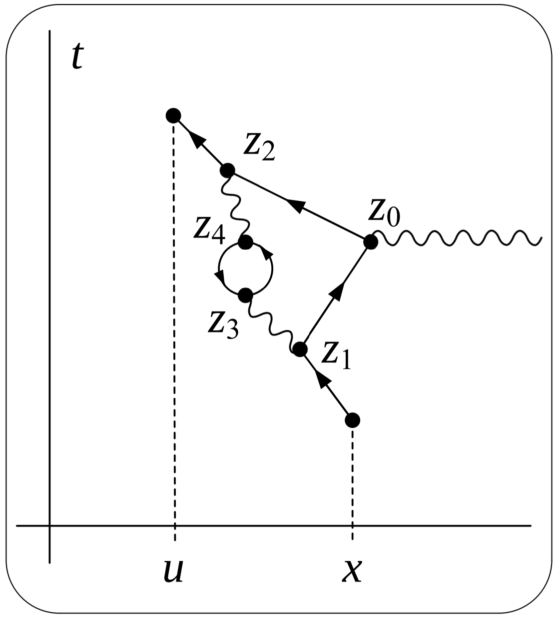


Figure 11: A possibility of higher order for improving the magnetic moment.

where the three vertices represent internal possibilities. The integral represents the fact that the vertices can be anywhere in spacetime yielding infinitely many mutually exclusive possibilities. Hence, we have to add all related arrows, that is, we have to evaluate an integral. If we add the amplitude ϕ^2 to the value 1, then we obtain the value 1.00116.

Later, improved experiments produced more accurate values. So we should look for more complicated diagrams. At the next level we have diagrams with 5 vertices representing internal possibilities. One of them is displayed in Figure 11. The electron emits a photon at z_1, the photon decays into an electron-positron pair at z_3, annihilating each other to form a new photon at z_4, which is absorbed by an electron at z_2.

In 1949, Karplus and Kroll [70] calculated the amplitudes of the diagrams with 5 vertices. Physicists found several algebraic errors in their calculations eight years later. Fortunately, the corrections affected only the fifth decimal digit.

Actually, there are more distinct Feynman diagrams, and it is a nice exercise to find out the geometrical picture of other ones with 5 vertices.

As before, adding all amplitudes together yields the overall amplitude:

$$\phi = \phi_1 + \phi_2 + \phi_3 + ... \tag{125}$$

In 1983, the magnetic moment was further improved with about 70 additional diagrams containing 7 vertices to the theoretical number 1.0011596 5246. This value coincides with the experimental value except for the last two digits. This accuracy is the same as if someone would measure the distance of about 3000 miles from Los Angeles to New York to within one millimeter. It took twenty years to obtain this extra accuracy. Kinoshita has improved this result to 13 decimal digits with a calculation involving 891 distinct diagrams.

[70] Feynman [1985, page 117]

Just as Einstein and many others with ideas on a solid basis were refused, Feynman had great difficulties. On the Pocono Conference of 30 March to 2 April 1948, arranged by Oppenheimer, Schwinger, Harvard's "Wunderkind", presented almost the whole day his developments in QED without any diagrams. Afterward late in the day, Feynman started his talk with his diagrams, wrote down the related integrals, and removed the infinities. He had to withstand frequent interruptions from Bohr, Dirac, Teller, and others, and no one seemed to be able to understand him. Feynman left the conference frustrated, even depressed.

5 Measurement

We describe the strange *wave-particle duality* and its consequences. It leads to the *measurement problem* in quantum theory, namely, the problem whether and how the wave function collapses to a certain outcome. This questioning has pushed one of the most challenging and partially nebulous debates about reality and quantum theory.

In our approach to quantum mechanics, using the trinity of time and distinguishing between possibilities, internal possibilities, and outcomes, many experiments become unbelievably simple to understand and are far from being strange or weird. They are described in a simple way by our probability recipe.

In Sections 5.1, 5.2, and 5.3 we discuss measurements as usually presented in many textbooks. In Section 5.4 we describe these phenomena from our point of view. Finally in Section 5.5, we discuss one of the fundamental principles in physics, consistent with our daily experience, namely *causality*: Events always happen in a fixed order, they cannot happen in different orders simultaneously. In recent literature it is stated that causality is violated.

5.1 The Wave-Particle Dualism and Paradoxes

One of the most well-known challenging problems in quantum theory is:

- No physicist has ever succeeded in detecting any particle traveling in one or the other path of a Mach-Zehnder interferometer, while observing simultaneously an interference pattern.

Similarly, the same holds true for slit experiments: the interference pattern vanishes whenever it is known through which slit the particle passes[71]. What is the nature of this impossibility? Is it a technical problem in the experimental setups? Is it a problem of interpretation? Or is it a consequence of incomplete physical models?

With our probabilistic recipe these questions can be simply answered. Please try it. Below, however, we will consider this problem and related ones as described frequently in the literature and from the historical point of view.

Interference patterns are typical for waves. In contrast, classical single particles exhibit either a deterministic behavior or a classical probabilistic behavior without any interference or cancellations. Experiments show that interference is destroyed whenever "*which-path information*" is available. Hence, we arrive at the well-known *wave-particle duality*, a concept in quantum theory where every physical object may be described as a particle, but also as a wave; see our critical remarks in Section 3.5.

Historically, the wave-particle duality originated from earlier works of Planck 1900 and Einstein 1905. The contribution of Einstein was that energy of a light beam is transmitted by discontinuous quanta, and not in the accepted form of a continuous wave propagation. He assessed his idea as a heuristic assumption and submitted a draft to some peers. Not surprisingly, as all ideas stated on a solid basis, but disturbing the actual consensus, his concept was rejected. Until 1922 Einstein was alone in believing the existence of light-quanta, called photons. However, Louis de Broglie, 1923, extended Einstein's idea to massive particles. He argued: if a beam of light, represented as a continuous wave, exchanges discontinuous photons when interacting with an appropriate environment, then, conversely, any particle can interact with environments as a wave. Moreover, he gave a mass to the photon. This statement seems to be proved. The photon mass[72] is about 10^{-51}kg.

In 1926, de Broglie's ideas and Schrödinger's wave equation were known, and Max Born proposed a probabilistic version of the wave model. Born[73] claimed that his wave model satisfies an invisible causality:

> *The particles motion obeys the probability laws, but probability itself propagates according to causality laws. [...] The question of knowing if the waves are something real, or a mere function for describing phenomena in an easy way, is a personal question. Personally, I like to think that a probability wave, even in a 3N-dimension*

[71]See also the discussions and examples in QUITE, Section 2.

[72]See for example Evans List: Photon mass and ECE Theory

[73]M. Born, Z..Phys. 437, 863-867 (1926)

> *space, is real, is something more than a tool for mathematical predictions. [...] For how could we believe probabilistic predictions, if by this notion we did not refer to something real and objective.*
> Born 1926

Thus, Born considered himself as a realist.

Bohr, 1927, introduced the concept of *complementarity* which states that a physical event cannot be understood in a single picture, but must be considered as complementary, that is, two pictures are absolutely necessary to obtain a complete information about a single microscopic object. Wave or particle as well as position or velocity are complementary quantities. The complementarity principle was mathematically justified a few months later by Heisenberg's uncertainty principle[74].

Notice that the three-dimensional position space and time are not complementary, they don't fulfill an uncertainty relation, but form the basis of physics in terms of the non-Euclidean space-time. This is doubtful. In QUITE, we have argued that the fundamental concepts of physics should be defined in terms of a six-dimensional Euclidean position-velocity space, and in fact this makes many things much easier.

Finally, we mention again a serious difficulty of the wave picture, if systems with N particles are considered. Obviously, *Schrödinger's wave equation* can be no longer an ordinary wave propagating in space-time. Instead, it propagates in the so-called configuration space of dimension $3N$. Even, for a small macroscopic system this dimension becomes astronomical huge.

[74]See QUITE, Section 4.17 for another interpretation of Heisenberg's uncertainty.

Figure 12: My experimental set-up costs about 90 Euros. The left half shows four polaroid filters, and on the right half there are two calcite crystals. The following experiments can be performed easily with these optical elements.

5.2 Polarization of Light

In order to understand the wave-particle complementarity more clearly, we consider linear polarization of light[75]. In the first part of this section, several experiments are considered and described with explanations that can be found in many textbooks. Only at the end of this section, we discuss these examples by using our probabilistic approach.

If a photon has passed a polarizer with transmission axis α, we write $|\alpha\rangle$ for this base state. Usually it is said that the photon is linearly polarized at this angle. As in Section 3.3, we assign to any transition from one initial state $|\alpha\rangle$ to another state $|\beta\rangle$ a complex *probability amplitude* (arrow in the complex plane):

$$\varphi_{\alpha\to\beta} = \langle\beta|\alpha\rangle \in \mathbb{C}. \tag{126}$$

The amplitude $\varphi_{\alpha\to\beta}$ is written as a bracket, the initial state $|\alpha\rangle$ is on the right hand side of the bracket, and the final state $|\beta\rangle$ is on the left hand side.

The following polarization experiments can be comprehended with optical elements such as with polaroid filters and calcite crystals. They are not expensive, see Figure 12.

Birefringent plates like calcite crystals split incident light with respect to their optical axis into a beam of vertically polarized light and a beam of horizontally polarized light. Both beams can be seen in the double pictures under the calcite crystals in Figure 12. In a more sophisticated experiment, where only one photon is interacting with the calcite, either detector D_x or D_y clicks, see Figure 13. Of course, we assume that the experimental setup is perfect,

[75]See also the discussions and further examples in QUITE, Section 2.

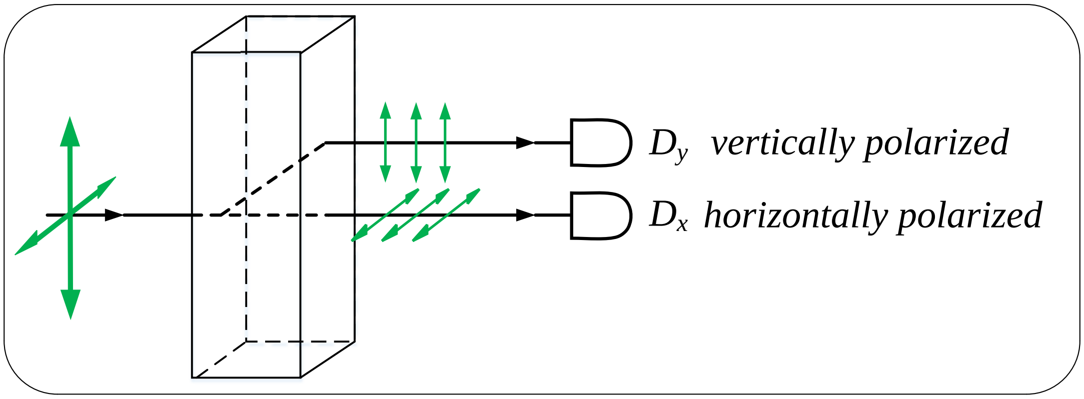


Figure 13: The birefringent plate splits incident light with respect to their optical axis into vertically polarized light and horizontally polarized light. If only one photon is in the experiment, either detector D_x or D_y clicks.

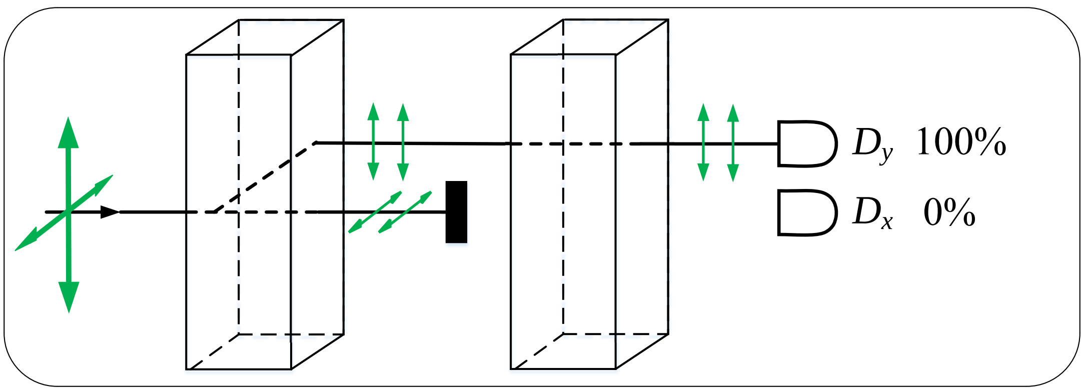


Figure 14: For two birefringent plates, where the horizontally polarized light after the first plate is blocked, only the vertical detector D_y clicks. If we block the vertically polarized light after the first plate, then only the horizontal detector D_x clicks. If the optical axis of the second birefringent plate is chosen opposite to the axis of the first plate, then we obtain analogous results, only the detectors change, that is, if the horizontally polarized light after the first plate is blocked, only the horizontal detector D_x clicks.

that is, the detectors are perfect, the single photon is not absorbed, and each emission of a photon is kept track of. Moreover, we assume that the incoming light is unpolarized.

Then what happens for unpolarized light is that in half of the cases the photon turns out to be vertically polarized, and in half of the cases the photon is horizontally polarized. This exhibits the particle-like picture of the photon. The photon has either chosen the upper beam with classical probability $1/2$, or the lower beam with the same probability.

The particle-like behavior of photons can be seen in experiment Figure 14. When the lower beam after the first plate is blocked, only the vertical detector D_y clicks. If we block the upper beam after the first plate, then only the horizontal detector D_x clicks. If the optical axis of the second birefringent plate is chosen opposite to the axis of the first plate, then we obtain analogous

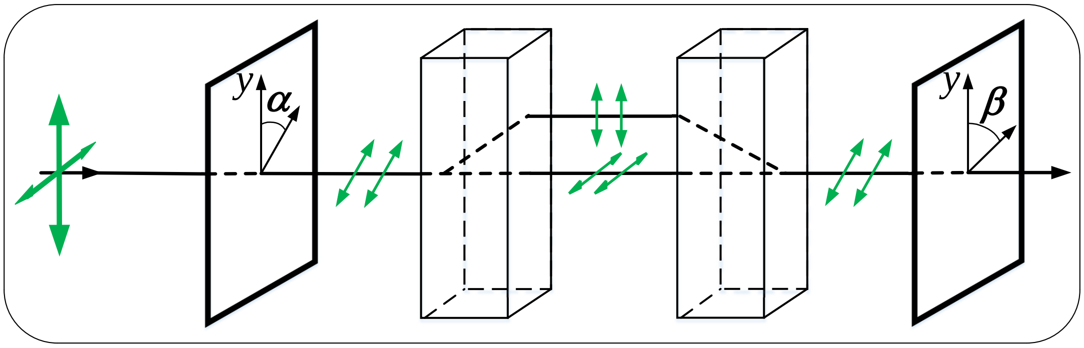


Figure 15: The first polarizer generates photons polarized at an angle α. The first birefringent plate splits into two beams of horizontally x-polarized and vertically y-polarized photons. These are recombined in a second birefringent plate which has an optical axis opposite to the first plate. According to the law of Malus the transition probability after the second polaroid is $\cos^2(\beta-\alpha)$.

results, only the detectors change. In other words, if the lower beam after the first plate is blocked, only the horizontal detector D_y clicks. Clicks of a positioned detector show and support the particle-like behavior of light and not wave-like pictures. If the photon passes the first plate, and the second plate has the same optical axis, thus both plates are identical, then the photon passes through the same beam of the second plate. In other words, if we repeat a measurement, then the experimental result does not change. If the optical axis of the second plate is chosen opposite to the axis of the first plate, then the photon passes the second plate on the opposite beam.

Let us now turn to the experiment displayed in Figure 15. We will show that a particle-like behavior of light is not sufficient to explain this experiment. There, the first polarizer, a polaroid filter with optical axis α, generates a photon polarized at an angle α. As before, the first calcite with optical axis at angle 0 implies a split into two beams of horizontally x-polarized and vertically y-polarized photons. The second calcite has an optical axis opposite to the first one. Experimentally, it turns out that these two paths are recombined as displayed in Figure 15. According to the law of Malus the transition probability after the second polaroid filter is $\cos^2(\beta - \alpha)$. In particular, if $\beta = \alpha$, then the photon passes the second polaroid filter with probability one.

Obviously, this cannot be explained with a particle-like picture. Let us consider the case $\beta = \alpha = \pi/4$. When the photon has passed the second calcite, then it is either horizontally polarized at angle 0 or vertically polarized at angle $\pi/2$. From the law of Malus it follows that only one half of the photons pass the second polaroid filter. This is not what happens. It contradicts the experimental results, namely, that in this experiment each photon passes the second filter. This type of experiments has been performed countless times, and they all state that with probability one the photon passes the second filter.

Since the particle picture is completely incompatible with these experimental results, in the literature it is stated that a photon cannot be a particle. Hence, lets think of a photon as a wave. Then after passing the first polaroid filter, the wave interacts with the first calcite and splits into two parts, one component in the upper beam, the other one in the lower beam. Both parts arrive at the second calcite with opposite optical axes. Constructive interference recombines both parts to the original wave, which then passes the second filter. Thus, this wave-like interpretation works very well for this specific experimental setup.

In summary, these experiments suggest to think of a photon as an oscillatory wave-packet that is constrained within a small region. Then on a large region the photon can be viewed as a localized disturbance that resembles a particle. Unfortunately, this nice picture of a particle as a localized wave-packet cannot describe our previous experiment in Figure 13. The calcite would split the wave-packet into two parts. The experiment says that exactly one of both detectors clicks. Hence, one part is responsible for the click, whereas the other part does nothing. All wave-packets are prepared in the same way, but sometimes we obtain horizontal polarization and sometimes vertical polarization, each with a chance of 1/2. But then there must be the chance 1/4 that both detectors don't click, and there must be a chance of 1/4 that both detectors click. This, obviously, doesn't happen. Hence, the photon is neither a classical particle nor a wave, nor a localized wave-packet. From where does the photon, viewed as a wave, know that one of its parts leads to a click in a detector, and so the other part is not allowed to interact with the other detector? The photon is not like a sound wave, or a water wave, or a local disturbance in any medium.

We close this section with Penrose's point of view[76]:

> *The quantum behaviour of things is indeed much more subtle than this. The wave that describes a quantum particle is not like a water wave or a sound wave, which would describe some kind of* **local** *disturbance in an ambient medium, so that the effect that one part of the wave might have on a detector in one region is independent of the effect that another part of the wave might have on a detector in some distant region. We see from experiment 1 [A Mach-Zehnder interferometer experiment] that the wave picture of a single photon, after it has been "split" into two simultaneous separated beams by a beam splitter, still represents just a* **single** *particle, despite this separation. The wave appears to describe some kind of* **probability distribution** *for finding the particle in various different places. This is getting somewhat closer to a description of what the wave is actually doing and, indeed, some people refer to such a wave as a* **probability wave**. *This, however, is not a satisfactory picture, because probabilities, being always positive quantities (or zero) cannot cancel one another out, as would be needed for an explanation of the absence of any responses at E in experiment 2.*
>
> *Sometimes people attempt a probability-wave type of explanation of this nature by allowing the probabilities somehow to become* **negative** *in places, so that cancellation can then take place. However, this is not really how quantum theory operates (see figure 2-5).*

[76]See page 136, Penrose [2016]

> *Instead it goes one step further than this by allowing the wave amplitudes to be* **complex numbers** *[...] These complex numbers are crucial to the whole structure of quantum mechanics.* Penrose 2016

In these notes we have introduced a new general concept of probabilities. The experimental setups solely allow us to distinguish between "as well as" possibilities (future things that are allowed in the experimental set) and "either-or" outcomes or elementary events (things that might happen when performing an experiment in the present). Each outcome can become a fact in the past. Outcomes are well-defined possibilities, but, in general, not every possibility is an outcome. Internal possibilities are responsible for interference. Complex probability amplitudes are assigned to possibilities, and via Born's rule probabilities are assigned to outcomes. The probability amplitudes for disjoint possibilities are added, and the probability for independent possibilities are multiplied. In the same way, the classical probabilities for disjoint events are added, and the probabilities for independent events are multiplied.

The experimental setup alone (without performing the experiment) determines the probability amplitudes and probabilities. The particles, during the execution in the present, obey the calculated probabilities. It is not necessary to attribute further properties to the particles, such as waves or points, or polarization, or spin, and so on, except the property that they can interact with the given machines or experimental setups.

We assign possibility amplitudes to the possibilities between machines, only. The amplitude $\langle \frac{\pi}{2} | \frac{\pi}{4} \rangle$, for example, where in the present a photon passes a first polarizer in base state $| \frac{\pi}{4} \rangle$ and then the next polarizer in base state $| \frac{\pi}{2} \rangle$ is a complex number which depends only on both polarizers, not on the photon itself. In the future interactions don't happen.

In our approach, the phrase "If a photon has passed a polarizer with transmission axis $|\alpha\rangle$, the photon is linearly polarized at this angle" means: A polarizer with any transmission axis, not the photon, is characterized by two mutually exclusive possibilities, called horizontal and vertical polarization. In the present, a photon interacting with this polarizer chooses exactly one of both possibilities.

Keep in mind: A photon is neither a classical particle, nor a wave like a sound or a water wave, nor a localized wave-packet, that is, a local disturbance in any medium. But the photon can interact with experimental setups, which may consist of various optical elements. These machines obey the probability theory as described above. This theory is based on complex probability amplitudes and the distinction between possibilities, internal possibilities and outcomes.

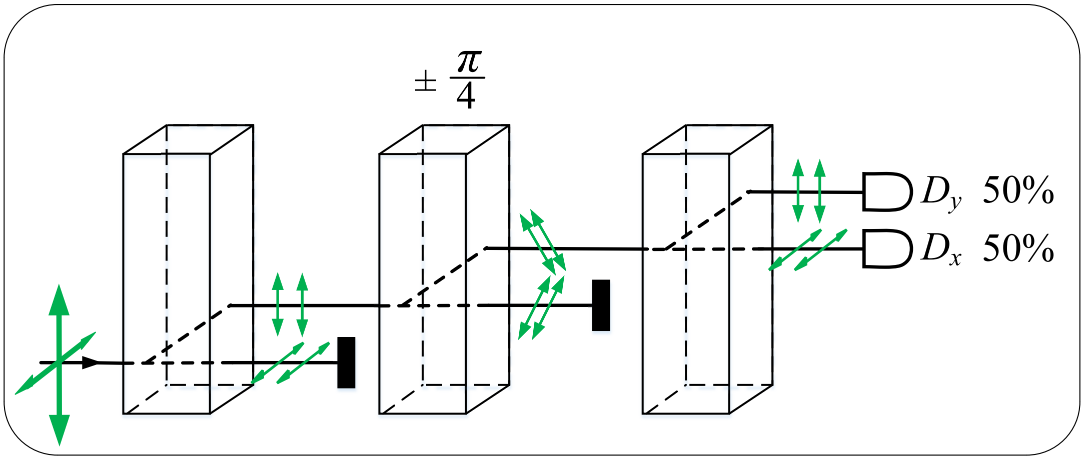


Figure 16: In this experimental setup the horizontally polarized photons $|0\rangle$ are blocked after the first birefringent plate. Then the photons with polarization $|-\pi/4\rangle$ are blocked after the second birefringent plate. Finally, after the third plate their polarization is measured with two detectors. Obviously, they have lost their original polarization. But if we replace the second plate by two plates that recombine both beams, as shown above, then all photons have the original polarization, that is, only the upper detector clicks.

5.3 The Measurement Problem

The wave-particle duality leads to the *measurement problem* in quantum theory, that is, the problem whether and how the wave function collapses to a certain event or outcome. This questioning has pushed one of the most challenging and partially nebulous debates about reality and quantum theory.

In this section we consider the widely used *von Neumann measurement scheme*. We present this scheme by using the customary language of quantum mechanics. At first, we look at polarization experiments with photons as discussed above. Later, we consider the general case introduced by von Neumann 1932.

Above we have discussed several experiments with photons, among them the experiment where both beams are recombined, see Figure 15. In this case it is not possible to know which path is used by a photon. We have interference due to internal possibilities, since only the upper detector clicks. Hence, the wave function is not reduced, but is unchanged equal to $|\pi/4\rangle$. In the experiment described in Figure 16, where the block, which recombines both beams, is simply replaced by one plate, the original polarization is lost.

In quantum theory, measurements are usually described by a microscopic system or any particle that interacts with the apparatus. The apparatus is equipped with detectors. The particle is described in the form of states.

Any polarization experiment equipped with two detectors at the right hand side, as in Figure 16, can be in three possible distinguishable base states: the state $|d\rangle$ before detecting a photon, the state $|d_y\rangle$ when a photon is detected in the upper detector, and the state $|d_x\rangle$ when a photon is detected in the lower detector. Moreover, we have a pointer that displays $0, +1$, or -1 for the base states $|d\rangle$, $|d_x\rangle$, or $|d_y\rangle$, respectively. In other words, the pairs $(0, |d\rangle)$,

$(+1, |d_x\rangle)$, and $(-1, |d_y\rangle)$ are the eigenpairs of an observable D that describes the apparatus.

Let the photon be horizontally polarized in base state $|0\rangle$, or vertically polarized in base state $|\pi/2\rangle$. The photon and the apparatus are composed via the tensor product of six mutually exclusive base states

$$|0\rangle|d\rangle, |0\rangle|d_x\rangle, |0\rangle|d_y\rangle, |\pi/2\rangle|d\rangle, |\pi/2\rangle|d_x\rangle, |\pi/2\rangle|d_y\rangle. \tag{127}$$

Now we look at the experiment in Figure 16, but we remove the block after the first plate such that all incoming photons pass the first plate. Moreover, we replace the second plate by two plates that recombine both beams as described in Figure 15. Obviously, each photon passes this experiment and is detected in one of the detectors.

If the incoming photon is horizontally polarized $|0\rangle$, the wave function of the whole system before detection is the tensor product state

$$|\xi\rangle = |0\rangle|d\rangle. \tag{128}$$

After the particle has passed the apparatus, it is detected in the lower detector, and the total wave function is

$$|\xi_x\rangle = |0\rangle|d_x\rangle. \tag{129}$$

Hence, this interaction changes the apparatus to base state $|d_x\rangle$, but leaves the polarization of the photon unchanged.

If the polarization of the incoming photon is vertically polarized $|\pi/2\rangle$, the wave function of the whole system before detection is the tensor product

$$|\xi\rangle = |\pi/2\rangle|d\rangle. \tag{130}$$

After the particle has passed the apparatus, it is detected in the upper detector, and the total wave function is

$$|\xi_y\rangle = |\pi/2\rangle|d_y\rangle. \tag{131}$$

As before, this interaction changes the apparatus to base state $|d_y\rangle$, but leaves the polarization of the photon unchanged.

Now we discuss the case where the incoming particle is in state

$$|\pi/4\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |\pi/2\rangle). \tag{132}$$

This is a base state corresponding to a polarizer with transmission axis $\pi/4$, according to the law of Malus. Then the wave function of the whole system before detection is the tensor product

$$|\xi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |\pi/2\rangle)|d\rangle = \frac{1}{\sqrt{2}}(|0\rangle|d\rangle + |\pi/2\rangle|d\rangle). \tag{133}$$

This initial state is a product state, thus not entangled. As we know, the whole system develops unitarily, thus linearly. Moreover, we know from the equations above how each single term evolves, namely each part of the sum on the right hand side of (133) develops as in (129) and (131). Therefore, linearity necessarily implies that the total wave function when the photon interacts with the detectors is

$$|\xi_{final}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|d_x\rangle + |\pi/2\rangle|d_y\rangle). \tag{134}$$

This state entangles the photon with the polarizing apparatus. According to Born's rule the probability that the apparatus shows +1 or −1 is 1/2.

Now we have some trouble with this measurement theory, where the complete system is described quantum mechanically. We know that the state of the detectors are either $|d\rangle$, $|d_x\rangle$, or $|d_y\rangle$. Paradoxically, these three mutually exclusive alternatives contradict equation (134). What does this mean for both detectors? Well, it is a quantum mechanical superposition which is usually interpreted that at the same time both detectors interact with the photon. One way to get out of this trouble is to measure both detectors with some second apparatus. After this second measurement, we expect that the wave function $|\xi_{final}\rangle$ is reduced to one part in the sum. But now, the additional second apparatus is part of a larger system that contains the second apparatus itself. Hence, we have to use a third apparatus, and so on. Obviously, in this infinite sequence of measurements we don't find a reduction of the wave function. Consequently, in quantum theory there seems to be nothing that converts probabilities into facts. To avoid this apparent paradox between our experience and the quantum measurement formalism, von Neumann contemplated a *collapse* that should select exactly one term in the superposition (134) with probability 1/2, perhaps caused by the consciousness of a human observer. This idea was later propagated by Wigner. In this connection one speaks today also of subjective theories.

The conclusions in the book of Susskind and Friedman[77] are:

> *Does the last entity to look at the system collapse the wave function, or does it just get entangled. Or is there a last looker? I won't try to answer these questions, but what should be apparent is that quantum mechanics is a consistent calculus of probabilities for a certain kind of experiment involving a system and an apparatus. We use it, and it works, but when we try to ask questions about the underlying "reality", we get confused.* Susskind and Friedman 2013

Below, we show that an underlying "reality" can be recovered without getting confused when using internal possibilities. Then quantum superpositions do not carry over to an infinite sequence of measurements.

This measurement paradox was illustrated by Schrödinger 1935, with the help of a cat, the famous and widely discussed *Schrödinger cat*. He described a Gedankenexperiment in which one could create a *superposition* of a macroscopic system. The cat's life or death depends on the state of a radioactive

[77] Susskind [2014], page 223

atom. If the atom has emitted radiation, a gun shoots the cat. Thus, formula (134) says that the cat remains alive and dead at the same time, leading to a superposition of a macroscopic system. That seems to be really weird, and it is. Schrödinger introduced the idea of dead-and-alive cats to illustrate the absurdity of the quantum mechanical formalism. Today, however, many physicists regard the dead-and-alive cat as quite real. Often, they use Schrödinger's cat as a way to illustrate and compare the strength and the weakness of particular quantum interpretations.

As early as 1952, Schrödinger[78] argued that there is no reason to collapse. He argued that the collapse is absurd: one cannot control the wave function in two different ways, sometimes by the wave equation, but occasionally by interference with the observer, not controlled by the wave equation. His solution was that the wave function never collapses, since a choice between a superposition of states is not necessary, provided our world is large enough. In terms of his famous cat experiment, both wave functions, leading to the "dead cat" and the "live cat", are equally real, that is, the world contains two "parallel worlds" one with the cat dead, the other one with the cat alive. The key interpretation is that these two worlds have always existed: starting with two cats alive, the two worlds change if one of the cats dies. Both worlds have exactly the same histories until the experiment is carried out. Then one world has a dead cat, the other world has a living cat. In fact, there is no splitting of worlds, no creation of new worlds, and no measurement problem, since all alternatives already exist. Hence, this means that our reality can be viewed as a path through the universe containing all these alternatives or parallel worlds, similar to bifurcation models. Obviously, this interpretation is closely related to the "many worlds" interpretation. This interpretation was originally formulated by Everett, but 5 years later in 1957. Schrödinger stagnated in darkness until in the 1980s David Deutsch developed his many worlds interpretation that is almost equal to Schrödinger's work, perhaps without being aware of this paper.

Finally, we consider the general case of von Neumann's measurement theory. The microscopic system or particle is represented in a Hilbert space H_P containing an orthonormal basis, say $\{|p_i\rangle\}$ with index set $i \in I$. The particle can interact with a measurement apparatus described by an observable D. This observable is represented by the eigenpairs $(D_i, |d_i\rangle)$ and defined on H_A. These pairs should correspond to mutually exclusive pointer positions. Moreover, these pairs are assumed to correspond to the outcome of an interaction or measurement if the particle is exactly in the base state $|p_i\rangle$. At the beginning before the measurement, the apparatus is initially in a "ready position", say $|r\rangle$. The total system, namely the particle and the apparatus, is described with the tensor product space $H_P \otimes H_A$. In this space the total system evolves unitarily in the form

$$\left(\sum_i \psi_i |p_i\rangle\right) |r\rangle \xrightarrow{t} \sum_i \psi_i |p_i\rangle |d_i\rangle. \tag{135}$$

[78] Schrödinger, Erwin. "Are there quantum jumps? Part I." The British Journal for the Philosophy of science 3.10 (1952): 109-123.

This unitary evolution of the wave-function above is referred to as a *pre-measurement*, since the measurement itself has not been completed. This wave-function is an entangled superposition of the particle with the measurement apparatus. We obtain the same paradox as above, that is, without any additional physical process, such as a collapse of the wave-function, it is not clear how to get a definite pointer position after the measurement. This is the well-known *wave-function reduction problem*, sometimes also called the *problem of definite outcomes*. There is another problem called the *problem of the preferred basis*: the expansion (135) of the superposed wave-function is not uniquely defined in general, since it is not clear which basis shoud be used. Hence, the measured observable is not uniquely defined[79].

[79] Schlosshauer [2015]

5.4 Measurement and Possibilities

In our interpretation, based on the trinity of time, the possibilities and the related complex probability amplitudes are defined completely by the experimental setup. The experimental setup defines the base states and possibilities uniquely, as can be seen for example in all polarization experiments. These are notions of the future, where nothing happens. Interactions are part of the present only. It turns out that in this model all measurement paradoxes vanish. No last "looker" or "many worlds" or subjective interpretations are required.

We discuss the measurement problem for the following experiment, which consists of 5 optical "machines" connected in series, say S, A, B, C, and D. The source S produces photons that are polarized at an angle $|\alpha\rangle$. Hence, this is a one-possibility apparatus. The next three machines A, B, C are birefringent plates, the first one A produces two mutually exclusive beams, the horizontal one $|0\rangle$ and the vertical one $|\pi/2\rangle$. The machines B and C recombine each incoming beam. The last machine consists of two detectors yielding two possibilities $|d_x\rangle$ and $|d_y\rangle$. Thus, the *elementary possibilities* of our experiment consist of all sequences $|\alpha abcd\rangle$, yielding the *possibility space*

$$P = \{|\alpha abcd\rangle : \ |\alpha\rangle \in S, |a\rangle \in A, |b\rangle \in B, |c\rangle \in C, |d\rangle \in D\}\,. \tag{136}$$

The field F_P is defined as the set of all subsets of the possibility space. The outcomes of this experiment are two disjoint possibilities, namely that exactly one of the detectors clicks. All other possibilities are internal. Hence, the outcomes are defined by the *non-elementary possibilities*

$$|\alpha ABCd_x\rangle = \{\alpha abcd_x : \ |a\rangle \in A, |b\rangle \in B, |c\rangle \in C, \}\,, \tag{137}$$
$$|\alpha ABCd_y\rangle = \{\alpha abcd_y : \ |a\rangle \in A, |b\rangle \in B, |c\rangle \in C, \}\,. \tag{138}$$

In the following we write shortly $|d_x\rangle$ and $|d_y\rangle$ for both outcomes.

We can simplify this experiment, since the recombining pair BC acts as an identity. Hence, it is sufficient to consider the three optical elements SAD, the possibility space

$$P = \{|\alpha ad\rangle : \ |\alpha\rangle \in S, |a\rangle \in A, |d\rangle \in D\}\,, \tag{139}$$

and, since the possibilities of apparatus A are internal, the outcomes are

$$|d_x\rangle = |\alpha Ad_x\rangle = \{\alpha ad_x : \ |a\rangle \in A\}\,, \tag{140}$$
$$|d_y\rangle = |\alpha Ad_y\rangle = \{\alpha ad_y : \ |a\rangle \in A\}\,. \tag{141}$$

The probability amplitude for the outcome $|\alpha Ad_y\rangle$, where $\alpha = 0$, is according to the multiply-and-add rule

$$\begin{aligned} \langle d_y|0\rangle &= \langle d_y|0\rangle\langle 0|0\rangle + \langle d_y|\tfrac{\pi}{2}\rangle\langle\tfrac{\pi}{2}|0\rangle \\ &= 1\cdot 1 + 0\cdot 0 \\ &= 1. \end{aligned} \tag{142}$$

Here we have used the fact that horizontal and vertical polarization are represented by orthonormal base states.

The probability amplitude for the outcome $|\alpha A d_y\rangle$, where $\alpha = \frac{\pi}{2}$, is

$$\begin{aligned}\langle d_y|\tfrac{\pi}{2}\rangle &= \langle d_y|\tfrac{\pi}{2}\rangle\langle\tfrac{\pi}{2}|\tfrac{\pi}{2}\rangle + \langle d_y|\tfrac{\pi}{2}\rangle\langle\tfrac{\pi}{2}|0\rangle \\ &= 0\cdot 1 + 0\cdot 0 \\ &= 0.\end{aligned} \tag{143}$$

Analogously, we get the probability amplitudes $\langle d_x|0\rangle = 0$ and $\langle d_x|\frac{\pi}{2}\rangle = 1$ for the outcomes $|\alpha A d_x\rangle$, where $\alpha = 0$ and $\alpha = \frac{\pi}{2}$, respectively.

Finally, for $\alpha = \frac{\pi}{4}$, we get the probability amplitudes

$$\begin{aligned}\langle d_y|\tfrac{\pi}{4}\rangle &= \langle d_y|0\rangle\langle 0|\tfrac{\pi}{4}\rangle + \langle d_y|\tfrac{\pi}{2}\rangle\langle\tfrac{\pi}{2}|\tfrac{\pi}{4}\rangle \\ &= 1\cdot\tfrac{1}{\sqrt{2}} + 0\cdot\tfrac{1}{\sqrt{2}} \\ &= \tfrac{1}{\sqrt{2}},\end{aligned} \tag{144}$$

and

$$\begin{aligned}\langle d_x|\tfrac{\pi}{4}\rangle &= \langle d_x|0\rangle\langle 0|\tfrac{\pi}{4}\rangle + \langle d_x|\tfrac{\pi}{2}\rangle\langle\tfrac{\pi}{2}|\tfrac{\pi}{4}\rangle \\ &= 0\cdot\tfrac{1}{\sqrt{2}} + 1\cdot\tfrac{1}{\sqrt{2}} \\ &= \tfrac{1}{\sqrt{2}}.\end{aligned} \tag{145}$$

It is elaborate but instructive to go into detail through all calculations without our simplification.

The possibility amplitudes are assigned to the possibilities between machines, only. For example, the amplitude $\langle\frac{\pi}{2}|\frac{\pi}{4}\rangle$ is a complex number describing a future execution of the experiment in which a photon passes the first polarizer in base state $|\frac{\pi}{4}\rangle$ and then the next polarizer in base state $|\frac{\pi}{2}\rangle$. Nothing happens in the future. Probability amplitudes depend only on the experimental setup, but not on interactions in the present.

Thus, the argument that linearity necessarily entangles the photon with the polarizing apparatus as displayed in formula (134), that is,

$$|\xi_{final}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|d_x\rangle + |\pi/2\rangle|d_y\rangle). \tag{146}$$

does not occur in our theory, since photons are absent. All results are based on the properties of the machines. In our mathematical model, we have instead the equation

$$|\pi/4\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |\pi/2\rangle), \tag{147}$$

which says that the possibility $|\pi/4\rangle$ of a plate can be expressed as the superposition of the possibilties $|0\rangle$ and $|\pi/2\rangle$. There is no entanglement.

5.5 Causality

One of the fundamental principles in physics, consistent with our daily experience, is that of *causality*: Events always happen in a fixed order, that is, they cannot happen in different orders simultaneously. In recent literature[80], it is stated that causality should be banned, in particular because of the rules of quantum mechanics differing much from classical mechanics. These rules seem to imply that causality is violated.

An example is the *quantum switch* where two operations A and B are connected. Then we obtain two mutually exclusive possibilities: Either "B follows A" or "A follows B". Then it is argued that in quantum mechanics both possibilities can be superposed, leading to an indefinite causal order such that both cases occur simultaneously. In other words, in the same manner as a material object can be at different places at the same time, both causal cases exist at the same time, thus destroying causality.

Experimentally, the quantum switch can be realized as an optical setup with a control qubit $|\psi\rangle_c$ defined in terms of the photon's polarization. The operations A and B, viewed as "black box operations", are optical machines applied to a target qubit $|\phi\rangle_t$ defined in terms of the transverse spatial mode. The control bit determines the order in which both operations are applied to the target qubit. When the control bit is $|0\rangle_c$, then operation B follows A. When the control bit is $|1\rangle_c$, then operation A follows B. But when the control qubit is in the superposition

$$|\xi\rangle = \frac{1}{\sqrt{2}}|0\rangle_c + \frac{1}{\sqrt{2}}|1\rangle_c, \tag{148}$$

then the output state of the system is in the superposition

$$|\Psi\rangle = \frac{1}{\sqrt{2}}BA|\phi\rangle_t \otimes |0\rangle_c + \frac{1}{\sqrt{2}}AB|\phi\rangle_t \otimes |1\rangle_c, \tag{149}$$

because of the bilinearity of the tensor product. Seemingly, quantum theory tells us that the daily experienced causality is violated.

Not surprisingly in our approach causality is not violated. Why? In the experimental realization of the quantum switch there is a source producing randomized photons. They pass a first calcite crystal with a variable polarization axis. The second calcite has a polarization axis along the z-axis with the two base states $|0\rangle_c$ and $|1\rangle_c$. Finally, the two machines describing both cases "B follows A" and "A follows B" are implemented. According to the requirements we change the polarization axis of the first calcite such that only photons with the desired polarization $|0\rangle_c$, $|1\rangle_c$, or their superposition pass the second calcite.

This realization satisfies the mathematical description of the quantum switch. All probabilities can be calculated with our recipe. Since nothing happens in the future, causality cannot be violated. Starting the experiment with one photon, in the present exactly one route is selected with the calculated probability. Hence, causality is also not violated in the present, and

[80]see for example Goswami et al. [2018]

trivially not in the past. The reason is our interpretation of time as the trinity future, present, and past, and of superpositions.

6 Conclusions

Our daily sense experiences - future, present, and past together with causality - were shattered by physicists on the altar " FASHION, FAITH and FANTASY ". In QUITE and in this Supplement it is shown that our daily experiences and observations are consistent within

- a modified theory of relativity, based on a position-velocity space with another time conception, thus based not on spacetime;
- a causal quantum single-world theory avoiding well-known paradoxes and describing future events in a world of machines where nothing happens;
- a unified probability theory that carefully distinguishes between possibilities, internal possibilities and outcomes, applicable to classical experiments up to quantum electrodynamics.

A major goal of these notes was to develop physics from a completely different perspective and imagination than usual, perhaps justifying the word "free climbing“. Surprisingly, we obtained the common mathematical formalism, but with unusual interpretations.

Finally, we mention that no-go theorems give only scant hints, if any. One example is Hardy's paradox where logic proves that a specific experiment is not realizable, although it is. Another one is the recent paper written by Frauchiger and Renner: "Single-world interpretations of quantum theory cannot be self-consistent“.

7 Appendix C: Keep in Mind

Keep in mind: Calculate the probability for the outcomes in Laplace experiments by using the *multiply-and-add rule*, that is, the probabilities for disjoint events are added, and the probabilities for independent events are multiplied. This rule is universal, since it applies also to classical probability as formulated by Kolmogorov, and to quantum probability.

Keep in mind: When solving probabilistic problems it is necessary to know precisely the sample space. Then many erroneous conclusions can be avoided, as the letters to Marilyn vos Savant demonstrate.

Keep in mind: When solving probabilistic problems, a precisely defined sample space may be not sufficient. The "principle of indifference" may be violated. For obtaining numerical probabilities, the process or program how the outcomes of the sample space are constructed may be necessary.

Keep in mind: In general, under nonlinear transformations the type of distributions changes, and the principle of indifference does not apply.

Keep in mind: Bertrand's paradoxes have shown that the sample space is not sufficient for calculating probabilities. Further information about the experiment is necessary. This information depends on the geometry of the experimental set up. Moreover, we assume a time trinity that distinguishes between future possibilities, present random access of outcomes in terms of momentary decisions, and the facticity of the past in terms of facts. Facts are elements of the set of outcomes, the latter are contained in the set of all possibilities. Time trinity allows, in a very simple way, to describe precisely experiments. A *probability* is defined as a map from the set of all outcomes into the set of real numbers between zero and one, and is related to the present. A *probability amplitude* is defined as a map from the set of all possibilities, including *internal elementary possibilities*, into the set of complex numbers with magnitudes between zero and one, and is related to the future. Squaring the magnitude of probability amplitudes for outcomes gives the probabilities, according to Born's rule.

Keep in mind: The recipe for calculating probabilities:
Given an experimental setup:

1. Define the possibility space P, and the internal possibilities.
2. Define the sample space Ω of outcomes.
3. Calculate the probability amplitudes for the outcomes by using the *multiply-and-add rule*, that is, the probability amplitudes for disjoint possibilities are added (*superposition*), and the probability amplitudes for independent possibilities are multiplied.
4. Calculate the probabilities for the outcomes using Born's rule.
5. Calculate with Kolmogorov's rules the probabilities for the classical non-elementary events.

The possibility space P and the field of subsets F_P are defined similarly as in classical probability theory the sample space Ω and the related field of subsets of the sample space. Moreover, in quantum theory the *multiply-and-add rule* holds true for probability amplitudes as well. The essential difference is (i) that amplitudes are complex numbers, (ii) that possibilities and outcomes are different quantities, and (iii) that internal possibilities, responsible for interference, are essential. Quantum theory can be viewed as a calculus with complex numbers that delivers numerical probabilities for outcomes based on experimental setups. This calculus is not restricted to microscopic systems. In contrast, it is mainly based on macroscopic machines. Quantum theory and classical probability theory are not different probability theories, but complement one another. We speak of classical experiments, if internal possibilities are absent. **This recipe completes our formulation of probability theory and the fundamentals of quantum mechanics.** Feynman's path integral, one of the mathematical equivalent formulations of quantum mechanics, is an immediate consequence of this recipe. Experiments, classical or quantum ones, can be explained by using this recipe.

Keep in mind: The *superposition principle* in the interpretation of the trinity of time is the unspectacular property of expressing possibilities of one machine in terms of the possibilities of other machines. Consequently, **a quantum state is not a property of one or more particles, but instead represents the properties of an experimental setup.**

Keep in mind: The slit experiment in 2012 with the large phthalocyanine molecules shows: (i) a molecule is not a wave, (ii) supports clearly our probabilistic approach, (iii) the pictures of the molecules with the video camera show that a material object is not at different places at the same time, and (iv) it leaves many quantum interpretations at least doubtful.

Keep in mind: When using the concept "trinity of time" with distinguishing between possibilities, internal possibilities and outcomes, then last looker or other strange properties are not required.

Keep in mind: Quantum mechanics turns out to be a generalized probability theory of Laplace experiments using complex numbers, the largest field of reasonable numbers according to Hurwitz. It shows the unbelievable simplicity of quantum theory, not strange, and easy to teach already in school.

Keep in mind: A photon is neither a classical particle, nor a wave like a sound or a water wave, nor a localized wave-packet, that is, a local disturbance in any medium. But the photon can interact with experimental set ups, which may consist of various optical elements. These machines obey the probability theory as described above. This theory is based on complex probability amplitudes and the distinction between possibilities, internal possibilities and outcomes.

Index

References

F. Allhoff. *Philosophies of the Sciences: A Guide*, Wiley-Blackwell, 2010.

A. Aspect, P. Grangier, and G. Roger. *Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: a new violation of Bell's inequalities*, Physical review letters **49**(2):91–94, 1982.

R. Bach, D. Pope, S.-H. Liou, and H. Batelaan. *Controlled double-slit electron diffraction*, New Journal of Physics **15**(3): 033018, 2013.

A.O. Barut, and R. Razka. *On Non-Compact Groups. I. Classification of Non-Compact Real Simple Lie Groups and Groups Containing the Lorentz Group*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. Vol. 287. No. 1411. The Royal Society, 1965.

H.P. Beck-Bornholdt, and H.H. Dubben. *Der Hund der Eier legt*, Rohwohlt, 2004.

Belzasar. *Results of a double-slit-experiment performed by Dr. Tonomura showing the build-up of an interference pattern of single electrons. Numbers of electrons are 11 (a), 200 (b), 6000 (c), 40000 (d), 140000 (e).* `https://commons.wikimedia.org/wiki/File:Double-slit_experiment_results_Tanamura_2.jpg`, provided with kind permission of Dr. Tonomura, 2012.

J.L. Bertrand. *Calcul des probabilite's*, Gauthier-Villars, Paris, 1889.

J. Cham, and D. Whiteson. *We Have no Idea, A Guide to the Unknown Universe*, Riverhead Books, 2017.

T. Chivers. *Neuroscience, free will and determinism: 'I'm just a machine'*, The Telegraph, October 12, 2010.

A.J. Coleman, V.I. Yukalov. *Reduced Density Matrices*, Lecture Notes in Chemistry No. 72, Springer, 2000.

K. Conrad. *The Hurwitz theorem on sums of squares*, 2010. `http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/hurwitzlinear.pdf`.

R.P. Crease. *The most beautiful experiment*, Phys. World **15**(9): 19–20, 2002.

C. Egli. *Feynman Path Integrals in Quantum Mechanics*, Seminar course in theoretical physics at KTH, Stockholm, 2004.

A. Einstein, *Physics and Reality* , Journal of the Franklin Institute 221.3, 349-382, 1936.

R.P. Feynman, *Space-Time Approach to Non-Relativistic Quantum Mechanics* , Rev. of Mod. Phys., 20, **22**, 367, 1948.

R.P. Feynman. *QED: The Strange Theory of Light and Matter*, Princeton University Press, 1985.

D. Frauchiger, R. Renner. *Quantum theory cannot consistently describe the use of itself*, arXiv:1604.07422v2 [quant-ph] 5 Oct 2018 2018. .

K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, A. G. White. *Indefinite causal order in a quantum switch* , arXiv:1803.04302v2, 2018

D Granberg, T. A. Grown.. *The Monty Hall dilemma*, Personality and social psychology bulletin, 1995.

D. Griffiths. *Introduction to Elementary Particles*, Wiley-VCH, 2004.

A. Hajek, *Probability, logic, and probability logic* , The Blackwell guide to philosophical logic, 2001.

L. Hardy. *Quantum mechanics, local realistic theories, and Lorentz invariant realistic theories*, Phys. Rev. Lett. 68, 2981-2984, 1992.

S. Hawking. *Does God play dice*, http://www.hawking.org.uk/does-god-play-dice.html, 2015

S. Hossenfelder. *Lost in Math. How Beauty Leads Physics Astray*, Basic Books, New York, 2018.

W.T. Irvine. *Realization of Hardy's thought experiment with photons. Physical review letters*, Physical review letters, 95(3), 030401, 2005.

C. Jansson. *Quantum Information Theory for Engineers: An Interpretative Approach*, DOI: https://doi.org/10.15480/882.1441, URI: httb://tubdoc.tub.tuhh.de/handie/11420/1444, 2017.

K. Kumericki. *Feynman diagrams for beginners*, arXiv preprint arXiv:1602.04182, 2016.

F. Laloë. *Do we really understand quantum mechanics? Strange correlations, paradoxes, and theorems*, American Journal of Physics 69.6, 655-701, 2001.

P-S. Laplace.. *A philosophical essay on probabilities*, English Edition, Dover Publications Inc., 1951.

H. C. Ohanian. *What is spin*, Am. J. Phys. 54(6), June 1986, 1986.

R. Penrose. *Fashion, Faith and Fantasy*, Princeton University Press, 2016.

T. Rudas. *Handbook of Probability*, SAGE Publications, 2008.

M. Schlosshauer. *Decoherence, the measurement problem, and interpretations of quantum mechanics*, Reviews of Modern physics 76.4 , 2005.

E. Schrödinger. *Are there quantum jumps? Part I*, The British Journal for the Philosophy of science 3.10 : 109-123, 1952.

G. Soff. *Quantenfeldtheorie. Lecture*, Technical University Dresden, 2002.

L. Susskind, and A. Friedman. *Quantum Mechanics: The Theoretical Minimum*, Basic Books, 2014.

A. Vazsonyi. *Which door has the Cadillac*, Decision Line 30.1, 19–20, 1999.

C.F. Weizsäcker. *Aufbau der Physik*, Carl Hanser Verlag, Munich, 1988.

C.F. von Weizsäcker. *Zeit und Wissen*, Carl Hanser Verlag, 1992.

C.F. Weizsäcker. *The Structure of Physics (Original 1985)*, T. Grönitz, and H. Lyre (eds.), Springer Netherlands, 2006.

S. Weinberg, *The trouble with quantum mechanics* , New York Review of Books 64.1, 51-53, 2017.