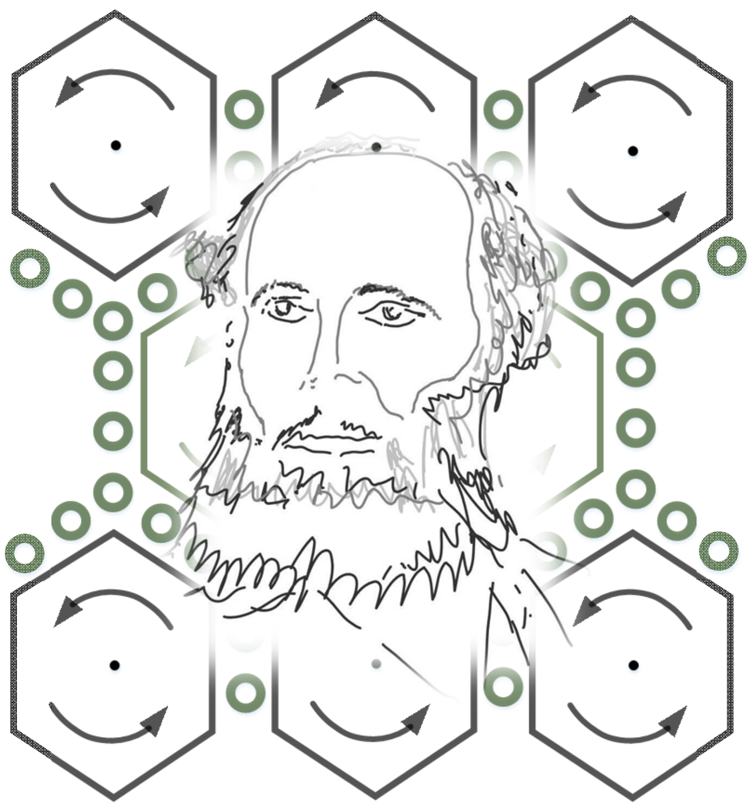
A Unified Treatment of Classical Probability, Thermodynamics, and Quantum Information Theory

Christian Jansson

September 14, 2021



https://doi.org/10.15480/882.3770

http://hdl.handle.net/11420/10330

Contents

Abstract

The major goal of these notes is an elaborate presentation of a probabilistic framework. This framework allows a formulation of classical probability theory, thermodynamics, and quantum probability with a common set of four principles or axioms. In particular, it provides a general prognostic algorithm for computing probabilities about future events. Our principles distinguish strictly between possibilities and outcomes. A well-defined possibility space and a sample space of outcomes resolves well-known paradoxes, and make quantum interpretations like "many worlds" or "many minds" superfluous. In addition, the superposition principle and the entanglement of systems obtain a new meaning from our point of view.

This framework offers an axiomatic approach to probability in the sense of Hilbert. He asked for treating probability axiomatically in his sixth of the twenty-three open problems presented to the International Congress of Mathematicians in Paris in 1900. We have applied our framework to various problems, including classical problems, statistical mechanics and thermodynamics, diffraction at multiple slits, light reflection, interferometer, delayed-choice experiments, and Hardy's Paradox.

Particular emphasis is also placed on C.F. von Weizsäcker's work, who developed his ur theory as early as the 1950s. Today, leading researchers continue his work under the name "Simons Collaboration on Quantum Fields, Gravity, and Information".

1 Preface

Sir Roger Penrose is a physicist, mathematician, philosopher of science, and Nobel Laureate in Physics in the year 2020. It should get the alarm bells ringing that this famous scientist, author of the excellent book "The Road to Reality, A complete Guide to the Laws of the Universe" said in an interview:

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

In 2010 he said farewell to our celebrated *"big-bang theory"* and proposed the old *"steady-state model"*:

> *The scheme that I am now arguing for here is indeed unorthodox, yet it is based on geometrical and physical ideas which are very soundly based. Although something entirely different, this proposal turns out to have strong echoes of the old steady-state model!* Penrose[1]

By the way, the widely glorified and seemingly experimentally verified message about the age of our universe would be wrong when believing Penrose.

Two recommendable, critical, recently published books are written by Cham and Whiteson[2] with the telling title "We Have no Idea, A Guide to the Unknown Universe", and by Hossenfelder[3] "Lost in Math. How Beauty

[1] Penrose [2010, Preface]
[2] Cham, Whiteson [2017]
[3] Hossenfelder [2018]

Leads Physics Astray“.

Perhaps, Penrose's first statement might be expressed in the form ” All physical models are wrong, but several are useful“.

This publication aims to provide a useful description of some new aspects of probability theory, thermodynamics, and quantum information theory, useful especially for students, engineers, and philosophers, but not exclusively. The mentioned new aspects rest on an evident categorization when describing and explaining experimental results. In particular, this categorization allows a useful probabilistic theory that is closely based on our daily experiences. It contains quantum mechanics without paradoxes and is teachable without magic. Moreover, we aim to answer the question: What has the second law of statistical thermodynamics and the reflection of one photon on a mirror in common? In particular, we show a close relationship between Feynman's path integral and thermodynamic multiplicity.

2 Introduction

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

More than 100 years ago, many scientists were uncontent with the missing clarity and rigor in probability theory since the basic concepts, such as randomness, events, or trials, turned out to be outside mathematics.

In 1900, Hilbert presented twenty-three fundamental problems. His sixth problem claimed to treat probability axiomatically, similar as in geometry. In particular, he called for a ”rigorous and satisfactory development of the method of average values in mathematical physics, especially in the kinetic theory of gases“. Many responses reemerged; see the excellently written paper of Shaver and Vovk[4].

In 1912 however, Poincaré[5] wrote

> *One can hardly give a satisfactory definition of probability.* H. Poincaré

Much later, von Weizsäcker[6] 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 2006

Even today, classical probability and its relationship to quantum probability are discussed somewhat nebulous. The right way how to assign probabilities to elementary events is a controversial philosophical discussion.

We shall investigate the following questions concerning probabilities:

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

[4] Shafer, Vovk [2006]
[5] Poincaré [1912] [Page 24]
[6] von Weizsäcker [2006, Page 59]

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, thermodynamics, and quantum probability?

The concept of probability is related to phenomena with uncertain outcomes or elementary events that form mutually exclusive alternatives. We can always distinguish between mutually exclusive events. They either happen or do not happen. But two or more elementary events cannot happen simultaneously.

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, probability describes a quantitative measure of the uncertainty of an outcome.

When investigating quantum probability, the debates and discussions become strange and weird. Fuchs[7] noted about the annual meetings and conferences:

> *What is the cause of this year-after-year sacrifice to the "great mystery?" Whatever it is, it cannot be for want of a self-ordained solution: Go to any meeting, and it is like being in a holy city in great tumult. You will find all the religions with all their priests pitted in holy war - the Bohmians[3], the Consistent Historians[4], the Transactionalists[5], the Spontaneous Collapseans[6], the Einselectionists[7], the Contextual Objectivists[8], the outright Everettics[9, 10], and many more beyond that. They all declare to see the light, the ultimate light. Each tells us that if we will accept their solution as our savior, then we too will see the light.* Fuchs 2002

Weinberg[8] 2017 writes in a worth reading 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 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

In these notes, we argue that probability theory, thermodynamics, and quantum probability can be formulated with a common set of rules or axioms, providing a predictive algorithm for computing probabilities about future events, like detector clicks. Our rules distinguish strictly between internal possibilities and outcomes. It is a theory characterizing the future and telling

[7] Fuchs [2002]
[8] Weinberg [2017]

us exactly what one should expect. Our approach avoids many well-known paradoxes and interpretations like ”many worlds“ or ”many minds“.

This article also contains and summarizes parts of two lecture notes[9], including some corrections. Both lecture notes contain much more issues, in particular, a new formulation of quantum mechanics.

In Section 3 we put things, such as mathematical or physical quantities, objects, or ideas into four different *categories.* This classification allows a better understanding of physics and probability. In particular, we replace the concept of an external time parameter with the *trinity* future, present, and past and show its consequences. We discuss the differences and relationships between possibilities, outcomes, and facts.

The primary goal of these notes is an elaborate presentation of a probabilistic framework consisting of four general principles which contain and marries classical probability and quantum probability. These principles form the content of Section 4. Readers only interested in probability theory can switch immediately to this section. It can be viewed as an axiomatic approach to probability in the sense of Hilbert, who asked for treating probability axiomatically in his sixth of the twenty-three open problems presented to the International Congress of Mathematicians in Paris in 1900. In particular, subsection 4.2 contains the central and fundamental part of these notes. In my lecture notes[10] this framework is applied to various problems, including classical problems, statistical mechanics, diffraction at multiple slits, light reflection, interferometer, delayed-choice experiments, and Hardy's Paradox.

In Section 5 we give a short survey about statistical thermodynamics and entropy, and we show its reconstruction. Perhaps thermodynamics is the most fundamental theory based on classical probability theory. Therefore, it is an essential touchstone for our probability theory. The basic ideas and tools of statistical thermodynamics are described. In particular, macrostates, microstates, multiplicities, and some examples are considered. Moreover, ”The Fundamental Assumption of Statistical Thermodynamics” and its relationship to the Boltzmann entropy and the second law of thermodynamics are discussed. Then a new form of entropy which we call *quantum Boltzmann entropy* is introduced.

Finally, in Section 6, we present a concise overview of quantum information theory, including some historical remarks and several aspects of von Weizsäckers fundamental work, the *ur theory.*

Acknowledgements I wish to thank Frerich Keil and Fritz Mayer-Lindenberg for their critical reading of parts of the manuscript, their feedback, and their suggestions.

Hamburg, Germany, September 2021
Christian Jansson

---

[9] Jansson [2017], Jansson [2019]

[10] Jansson [2019]

3 Categorization and Representation

Classical categorization dates back to Plato and Aristotle. They grouped objects according to their similar properties. In their understanding, categories should be clearly defined and mutually exclusive. Categorization schemes apply in language, prediction, decision making, types of interaction with the environment, and several other areas. Categories are the basic concepts of our thinking, and thus have a significant influence on all scientific descriptions.

A simple example in physics is the categorization of waves into longitudinal waves versus transverse waves versus surface waves, or electromagnetic waves versus mechanical waves versus quantum wave functions[11]. This type of categorization is derived from our observations and experiments for periodic vibrations.

We introduce a categorization of physics that is very close to our daily sense experiences. It is related to the four questions:

(1) What are the objects that have structure?

(2) What might happen in the future?

(2) What happens momentarily?

(4) What has happened?

A consequent application of these four categories to physics leads to surprising results, especially in probability theory, thermodynamics, and quantum information theory. It results in a different interpretation of the quantum *superposition principle*, which avoids the strange idea that a material object is in several places simultaneously.

It is not the intention to explain this world in an ontological sense or exhibit the basic structures of reality. But we would like to describe probability and physics in a useful way, preferably without paradoxes or magic.

3.1 Time and Structure

Even Neanderthals would immediately agree with the following observations: (1) The world is structured. There are buffalos, trees, and spears. (2) We do not know what will happen, for example, whether we will be successful on the next hunt. (3) There is no rest, and things happen momentarily. Right now, the spear hits the buffalo. (4) Many things have happened. Today we were successful on the hunt.

These four simple observations are so fundamental that physics perhaps should be described and understood in terms of these observations. Which notions and quantities belong to which observation?

Indeed, a primary goal of this publication is to describe physics in terms of these four categories. We classify things, such as mathematical or physical quantities, objects, or ideas in accordance with these *categories*.

[11]See ”The Physics Classroom > Physics Tutorial > Vibrations and Waves > Categories of Waves“ for more details.

(1) **Structure: Things are structured.** We receive all information from detector clicks, that is, from special machines in the broadest sense. Machines are best described or characterized by their possibilities. Possibilities form mutually exclusive alternatives. They either happen or do not happen, but two or more possibilities cannot happen simultaneously. Experimental set-ups consist of various machines which form a *web of relationships*. The set-up itself creates a machine. In Section 4, we argue that outcomes or elementary events consist of sets of possibilities.

In slit experiments, for example, the path from the source via any slit to a particular detector is a possibility. The outcome, where a particle is detected, consists of the set of paths from the source to this detector. Possibilities and outcomes belong to the structure of the experiment and should not be mixed up with probabilities or dynamics. The action, a geometrical functional which takes a path as its argument and has a real number as its result, and the related amplitudes are structural quantities. Moreover, in Section 5, we show that thermodynamic microstates correspond to possibilities, macrostates correspond to outcomes, and entropy and some versions of the second law of thermodynamics belong to the category structure. For example, the second law is sometimes formulated as[12]: Removing any constraints from an isolated thermodynamical system, thus changing the experimental set-up, will increase entropy. If the experimental set-up is not changed, then the entropy does not change.

(2) **Future: Things that might happen.** The future is characterized, in contrast to the structural category, as a timeless probabilistic framework. It is best described by the phrase "What might happen if nothing happens“. The future is prognostic and includes the principle of indifference, classical probability, and quantum probability.

(3) **Present: Things that happen momentarily.** *Dynamics* take place in the present: There is no rest (see also Section 4.8), and physical particles and systems tend to move towards states of larger probability. Events that are expected to occur more frequently occur more frequently. The motion can be thought of as a sequence of *collapses* in accordance with the probability distribution. In other words, the dynamics, say Wiener processes and zigzag Brownian motion, obey statistical concepts only, not classical deterministic laws like Newton's equations or Maxwell's equations. The latter equations approximate the stochastic dynamics under certain conditions and serve to calculate actions, which are required for calculating probability amplitudes.

(4) **Past: Things that have happened.** Relative frequencies, measurements, and occurred interactions belong to this category. These things form our history and usually change the structure. Experiments must first be built up before they can be carried out. This requires a lot of interaction.

[12] Ben-Naim [2018] Section 4.5

3.2 Trinity of Time

> *If I look at where we have paradoxes and what problems we have, in the end they always boil down to this notion of time.* Renato Renner[13]

Quantum theory, often referred to as the fundamental physical 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. This replacement is very close to our sense experiences and avoids many paradoxes. In this section, we present a short and rough overview. More details and several applications are considered and discussed in my lecture notes[14].

We consider quantum mechanics as a theory of probabilistic predictions that characterize the future only. The *future* is timeless, nothing happens. Quantum mechanics has to be understood prognostic. It is a probability theory that assigns to mutually exclusive alternatives, describing *possibilities* of machines, experimental set-ups, or physical systems, complex numbers which are called *probability amplitudes*.

We look 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 $k = 1, 2, 3, 4, 5, 6$. When we assign to each possibility the probability amplitude $1/\sqrt{6}$, then squaring according to Born's rule, gives the probability $1/6$.

Now, we consider the polarization experiment[15] in Figure 1. 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.

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,

[13] Wolchover [2020]
[14] Jansson [2017]
[15] For more details see Jansson [2017, Sections 2.3, 2.4, 2.5]

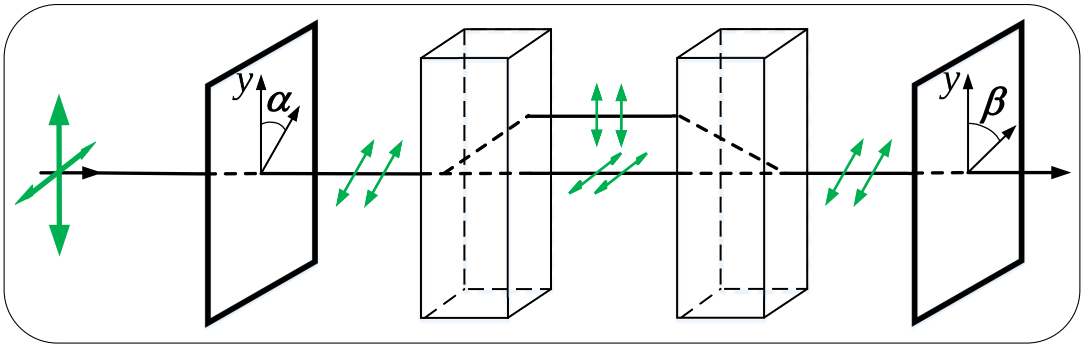


Figure 1: The first polarizer generates photons polarized at an angle α. The first birefringent plate splits into two beams of horizontally x-polarized and vertically y-polarized photons. These are recombined in a second birefringent plate which has an optical axis opposite to the first plate. According to the law of Malus, the transition probability after the second polaroid is $\cos^2(\beta - \alpha)$.

are called *outcomes* or *elementary events*. 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. They may consist of various internal alternatives, which we call *internal elementary possibilities*. We call physical models *classical* if all outcomes consist of precisely one elementary possibility.

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 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 amplitudes[16]. There are several paths through the slits, describing internal possibilities that lead to the same outcome. Thus, this is a non-classical model. However, if we position detectors at the slits, then we obtain a classical model.

Let us look at the outcomes for the polarization experiment in Figure 1:

- (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 the future and the present. In the literature, the property that there may be more possibilities

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

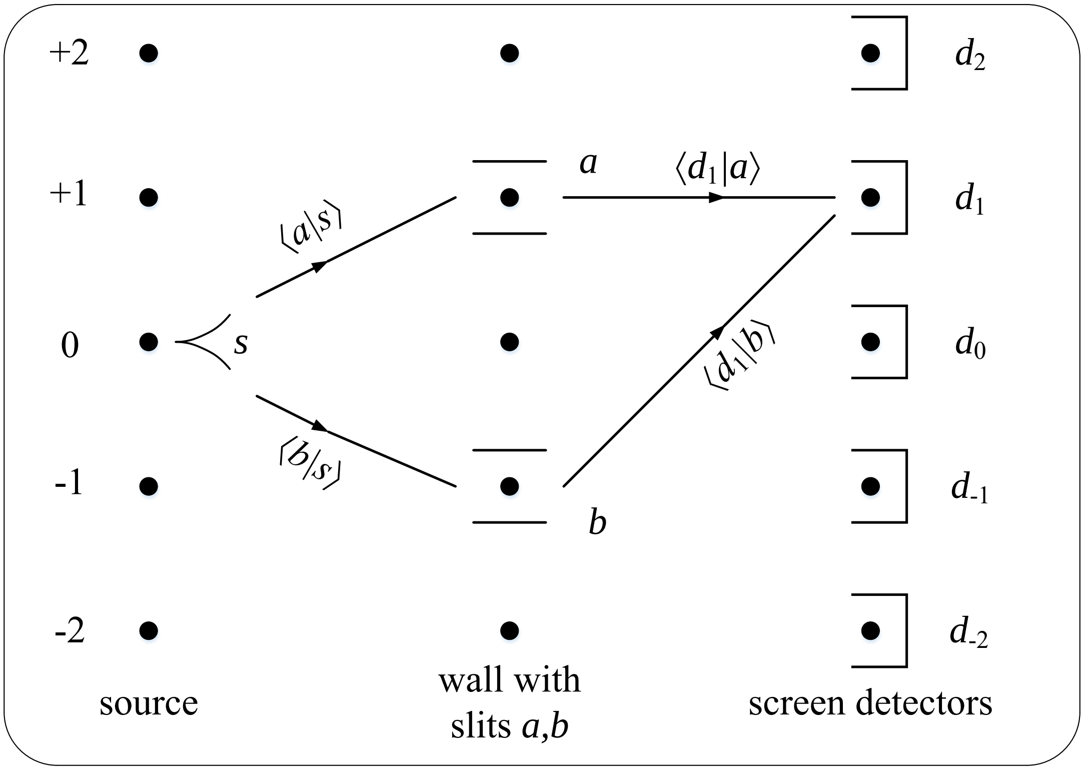


Figure 2: The double-slit experiment described for a discrete spacetime. The particle leaves source s, passes one of the two slits a or b, and is finally detected in d_1.

than outcomes leads to statements like " a material object occupies several locations simultaneously ". The failing distinction between past, present, and future in physics is the reason for many paradoxes in current quantum theory. In our categorization, the probabilities belong to the future where nothing happens. Only in the present, a material object chooses one elementary possibility in agreement with the probabilities. The object has the tendency to select possibilities with higher probabilities. However, occasionally the object might also choose possibilities with lower probabilities.

Deterministic models, like classical mechanics or electromagnetism, are described in terms of differential equations that don't allow alternative solutions provided initial conditions are given. 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. In general, quantum mechanics is non-classical since outcomes can be reached via several internal elementary possibilities. To summarize, 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 classical probabilities are non-negative numbers, cancellation or interference cannot occur. In contrast, a *probability amplitude* is defined as a map from the set of all possibilities into the set of complex numbers. Squaring the magnitude of probability amplitudes for outcomes 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. The concept of *relative frequencies* describes 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 via concepts of the past.

It is essential 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 precisely 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. In the past, this location becomes a fact.

This trinity of time is closely related 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. Time t appears in almost all physical 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

$$\frac{d}{dt}(m\dot{x}) - kx = 0. \tag{1}$$

The idea of equations without variable time seems questionable at first or even very strange. But after a while, we can 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*

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

is the conserved total energy, that is, the sum of kinetic and potential energy. This equation describes the harmonic oscillator just as well without time t, implicitly. It represents an ellipse in the phase space.

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 geometric quantities in their explicit form.

The fundamental theory of statistical thermodynamics, which can be applied to almost all physical models, independent of which concrete laws the systems satisfy, is timeless[17]. The entropy as well as the second law of thermodynamics has nothing to do with time. In Section 5 we reconstruct statistical thermodynamics with our probabilistic framework below.

In my lecture notes[18], several arguments are given to choose an Euclidean (3+3)-position-velocity space as a basis of physics, without any time parameter. It was shown how to reconstruct the mathematical formalism of special relativity by constructing clocks in this position-velocity space. In particular, we derived the key of relativity theory, namely the *Lorentz transform*, without any assumption about "propagation of light". Hence, Einstein's derivation of the relativistic spacetime can certainly be questioned.

Von Weizsäcker[19] emphasizes at various places the fundamental difference between the "factual past“ and the "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.

At a first glance, the presented trinity seems to create another time concept. However, this concept is completely different from other ideas about time since it rotates the past into the future, the future into the present, and the present into the past. Moreover, it differs significantly from the well-known "arrow of time“ which is discussed controversially. This thermodynamic arrow expresses a "one-way property of time“, and was created in 1928 by Eddington in his famous book "The Nature of the Physical World“. However, Ben-Naim[20] writes:

> *Reading through the entire book by Eddington, you will not find a single correct statement on the thermodynamic entropy.* Ben-Naim 2017

4 The Calculus of Probability Amplitudes

At the beginning of the twentieth century, mathematicians realized that probability theory seemed to use concepts outside mathematics like events, uncertainty, trial, randomness, probability. They were dissatisfied, and Hilbert asked for a clarification in his sixth of the twenty-three open problems presented to the International Congress of Mathematicians in Paris in 1900. He claimed to treat probability axiomatically. A nice presentation of the history of probability is presented by Shafer and Vovk[21]. In this section, we want to present a probabilistic framework consisting of four general principles

[17] Ben-Naim [2018]
[18] Jansson [2017, Sections 4.13 and 4.14]
[19] von Weizsäcker [1988], von Weizsäcker [1992], von Weizsäcker [2006]
[20] Ben-Naim [2018, p. 3]
[21] Shafer, Vovk [2006]

which unify classical probability and quantum probability. We show for various applications how these principles work. We have applied our probabilistic framework to classical problems, statistical mechanics, information theory and thermodynamics, double-slit and diffraction at multiple slits, light reflection, interferometer, delayed-choice experiments, and Hardy's Paradox; see my two lecture notes[22].

4.1 The Space of Possibilities

In physics, we observe or measure outcomes of experiments only. In the following, we investigate an imaginary experiment, say ABC, consisting of three machines described by finite sets A, B, and C, which are connected in series. The generalization to a large number of machines $A, B, C, D, \ldots.., K$ is straightforward. Our notation is close to Feynman's famous publication[23].

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* $a \in A, b \in B$, and $c \in C$. *Elementary* means that the possibilities cannot be further separated. *Mutually exclusive* means that the elementary possibilities are non-overlapping and distinguishable. In the present, a particle or a system interacts with the machines by choosing exactly one possibility, but two or more possibilities cannot be chosen simultaneously. For example, viewing space as a machine of positions, a single material object cannot occupy several locations simultaneously.

Possibilities of machines belong to the category structure. The elementary possibilities of the complete experiment ABC consist of all triples abc. Typically, such a triple means that, in a future interaction of a particle or a system with the experimental set-up, it starts by choosing a possibility $a \in A$, then interacts with B by choosing any possibility $b \in B$, and finally chooses an elementary possibility $c \in C$ where it is detected. We call the set of all *elementary possibilities* abc the *possibility space* ABC of the experiment, that is,

$$\mathbf{P} = ABC = \{p = abc : \ a \in A, b \in B, c \in C\}. \tag{3}$$

The experimental set-up itself can also be viewed as one single machine[24].

The set of all subsets of ABC, is denoted by $\mathbf{F}_{ABC}$. We identify the elementary possibilities abc with $\{abc\}$, the subsets consisting of one element.

[22] Jansson [2017], Jansson [2019]. These notes contain some bugs. This publication includes corrections and further developments.

[23] See Section 2, Feynman [1948]

[24] In this subsection, we assume that $\mathbf{P}$ is a finite set.

Other subsets are the *non-elementary possibilities*, such as

$$abC := \{abc : \ c \in C\}, \tag{4}$$
$$aBc := \{abc : \ b \in B\}, \tag{5}$$
$$Abc := \{abc : \ a \in A\}, \tag{6}$$
$$aBC := \{abc : \ b \in B \ c \in C\}, \tag{7}$$
$$AbC := \{abc : \ a \in A \ c \in C\}, \tag{8}$$
$$ABc := \{abc : \ a \in A \ b \in B\}, \tag{9}$$
$$. \tag{10}$$

For instance, the possibility aBc means that in a future interaction of a particle with the experimental set-up ABC, the particle has chosen the elementary possibility a, finally has chosen c, and further, it must have chosen some intermediate, not further specified, elementary possibility b of machine B. It may be that we are not interested in the possibilities of B. But it may also be that the interaction with B is unknown, and the experimental set-up does not allow the knowledge of a specific $b \in B$. In other words, b cannot be 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 defined explicitly. 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 represents 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 a possibility $a \in A$ in the present fortunately need not go through all internal possibilities $b \in B$ simultaneously, as it is usually assumed in quantum theory. Similarly, the possibility aBC means that, in the present, there is some interaction with A in a, but the interactions with B and C are not further specified. Hence, we can identify aBC with a itself. Now, we have defined non-elementary possibilities in terms of subsets of the possibility space. But what are outcomes? Well, this is the information given to the environment.

Let us consider three examples. 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 is not given outside since it is not detected. Hence, only the subsets $sWd_m \in \mathbf{F}_{SWD}$ define *outcomes*, and thus may become *facts* in the past.

The second one is the classic experiment where we throw a die two times. This can be viewed as three machines ABC, where machine A describes the first throw by the set of possibilities $\{1, 2, 3, 4, 5, 6\}$, B describes the second throw, and C describes the outcomes of both throws. There are no internal possibilities, and each elementary possibility, say abc, is an outcome and thus can become a fact. For example, $abc = 66(6,6)$ is the elementary possibility that the die would show 6 in each throw. The possibility space ABC coincides with the classical sample space $\mathbf{O}$ of outcomes[25]. Hence, we have a classical experiment. For a fair die, the probabilities for the outcomes are $1/6^2$.

Let us change this experiment such that the result of the second throw described by B cannot be recovered. In other words, the possibilities of B are internal. Then the outcomes are aBc and thus differ from the elementary possibilities abc. Clearly, a change of the experimental set-up changes the probabilities. For a fair die, the probabilities are $1/6$.

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

$$\mathbf{O} = \{F = aBc : \; a \in A, c \in C\}. \tag{11}$$

All other subsets of ABC are not outcomes. Notice that the outcomes are disjoint sets which partition the possibility space, that is,

$$ABC = \bigcup \{F \in \mathbf{O}\}. \tag{12}$$

The outcomes are characterized by their elementary possibilities $abc \in aBc$, which we call the *accessible elementary possibilities*.

Keep in mind that the notions of possibilities and outcomes are timeless and belong to the category structure, whereas the probabilities belong to the prognostic future.

4.2 A Unified Probabilistic Framework

After this physical motivation, we describe our probabilistic framework from the mathematical point of view. 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, probability describes a prognostic measure of the uncertainty of an outcome. It belongs to the category *future*. In contrast, the *relative frequency* belongs to the *past*, since it is defined as the number of experiments in which a specific outcome occurs divided by the number of experiments performed. It makes probability empirically testable, at least approximately.

An experiment is described by three sets:

(i) The *possibility space* $\mathbf{P}$ consisting of all elementary possibilities $p \in \mathbf{P}$.

(ii) The *possibility algebra* (also called *field*) $\mathbf{F}$ defined as the collection of subsets of the possibility space that contains $\mathbf{P}$ itself, and is closed under complement and under countable unions. The subsets $F \in \mathbf{F}$ which don't coincide with the elementary possibilities $\{p\}$ are called *non-elementary*.

[25]Based on the word "outcome", we use the letter $\mathbf{O}$ for the sample space and not, as usual, Ω. The latter is used in the Boltzmann entropy equation.

(iii) The sample space $\mathbf{O}$ of outcomes $F \in \mathbf{F}$ which form a partition of the possibility space, such that each elementary possibility $p \in \mathbf{P}$ is contained in exactly one outcome F.

These definitions belong to the category structure. In Section 4.1, the possibility space is $\mathbf{P} = ABC$, the field $\mathbf{F} = \mathbf{F}_{ABC}$ is the set of all subsets of ABC, and $\mathbf{O}$ is defined in (11).

Moreover, we assume:

(iv) A *probability amplitude*[26] is given, which is defined as a mapping φ from the field of possibilities $\mathbf{F}$ into the set of complex numbers:

$$F \to \varphi_F = \varphi(F) \in \mathbb{C}, \quad F \in \mathbf{F}. \tag{13}$$

We call the quadruplet $(\mathbf{P}, \mathbf{F}, \mathbf{O}, \varphi)$ *possibility measure space.* This space belongs to the category structure.

Motivation for the description of probabilistic and physical foundations with complex numbers can be found in Jansson[27]. See also the recent publication of Wood[28]. There, it is argued that complex numbers are fundamental and essential for describing reality. Notice that in the literature a measure is a non-negative function in contrast to amplitudes. We consider a measure with complex numbers.

The possibility measure space satisfies four general principles. The **first principle** states that for any countable set of pairwise disjoint possibilities $F_m \in \mathbf{F}$, such that $F = \cup_m F_m$, it is

$$\varphi_F = \varphi\left(\bigcup_m F_m\right) = \sum_m \varphi_{F_m}. \tag{14}$$

This rule is called the *superposition of probability amplitudes.* It expresses in a slightly different manner Feynman's first principle: "When an event can occur in several alternative ways, the probability amplitude for the event is the sum of the probability amplitudes for each way considered separately"[29]. Notice that Feynman does not distinguish between outcomes and possibilities.

The **second principle** is *Born's rule* which transforms the probability amplitudes of outcomes F to probabilities $\Pr(F)$:

$$\Pr(F) = |\varphi_F|^2 \text{ for all } F \in \mathbf{O}, \text{ and } \sum_{F \in \mathbf{O}} |\varphi_F|^2 = 1. \tag{15}$$

Thus, we obtain by computing the square of the magnitude of probability amplitudes the classical probabilities for the outcomes. If we sum up the probabilities of all outcomes, we get one. Hence, with Born's rule we obtain

[26] We use the notation in Feynman [1948] p.4
[27] Jansson [2017], Section 2.2
[28] Wood [2021]
[29] Feynman Lectures [1963] p.1-16

a probability measure on the sample space $\mathbf{O}$, and we can use Kolmogorov's rules for obtaining probabilities for the subsets of the sample space.

An experiment is called *deterministic* if the sets $\mathbf{P}$ and $\mathbf{O}$ consist of one element. In this case, Born's rule implies that the probability of the unique outcome is one. An experiment is called *classical* if both sets $\mathbf{P}$ and $\mathbf{O}$ coincide. In this case, classical probability theory applies from the very beginning.

(Consistency, $U(1)$ symmetry): Our probabilistic framework is consistent, that is, it does not lead to a contradiction. Moreover, all probabilistic statements are invariant if one transforms all elementary possibilities with one element of $U(1)$.

At first, we show that the probability amplitude is well-defined, that is, the amplitude φ_F does not depend on the partitioning of F. If F contains only one element, nothing is to proof. For two disjoint elements we have $F = \cup \{F_1, F_2\}$ and $\varphi_F = \varphi_{F_1} + \varphi_{F_2} = \varphi_{F_2} + \varphi_{F_1}$ is well-defined. If F is the union of three pairwise disjoint possibilities F_1, F_2, F_3, we can partition $F = \cup \{F_1, F_2, F_3\}$ as follows:

$$F_1, F_2, F_3;\ \cup \{F_1, F_2\}, F_3;\ \cup \{F_1, F_3\}, F_2;\ \cup \{F_2, F_3\}, F_1. \tag{16}$$

Since complex addition is associative and commutative, in all cases our first principle yields

$$\varphi_F = \varphi_{F_1} + \varphi_{F_2} + \varphi_{F_3}. \tag{17}$$

Hence, φ_F is well-defined. The same is true if F is partitioned into more than three elements:

$$\varphi_F = \sum_m \varphi_{F_m}. \tag{18}$$

The second principle requires that the sum of the square of the magnitudes of probability amplitudes for all outcomes is one. This is a simple normalization condition that can always be achieved.

Moreover, if we multiply all probability amplitudes with the same element $e^{i\phi} \in U(1)$, then due to Born's rule, the probabilities do not change.

The fundamental symmetry group $U(1)$ leaves the inner product of two complex numbers and thus their norm constant. This group is locally isomorphic to the symmetry group $SO(2)$ of rotations in a two-dimensional real space. $U(1)$ gauge symmetry is well-known in quantum electrodynamics, where one cannot measure the absolute phase of the wave functions of electrons and photons.

As the most simple example, consider a fair coin toss. The two elementary possibilities are Heads H and Tails T. They define the possibility space $\mathbf{P} = \{H, T\}$. The field of possibilities is

$$\mathbf{F} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}. \tag{19}$$

The set of outcomes coincides with the two elementary possibilities:

$$\mathbf{O} = \{\{H\}, \{T\}\}, \tag{20}$$

They form a partitioning of the possibility space. We define

$$\varphi_\emptyset = 0,\ \varphi_{\{H\}} = \frac{1}{\sqrt{2}},\ \varphi_{\{T\}} = \frac{1}{\sqrt{2}}. \tag{21}$$

Then our first principle yields

$$\varphi_{\{H,T\}} = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \sqrt{2}. \tag{22}$$

The fact that $|\varphi_{\{H,T\}}|^2 \neq 1$ does not imply any contradiction, since Born's rule defines probabilities only for the outcomes.

It is also simple to model an unfair coin, say with probability 1/3 for Heads and 2/3 for Tails. In this case we define the possibility space in the form $\mathbf{P} = \{H, T_1, T_2\}$, where the outcome Tails is the set $\{T\} = \{T_1, T_2\}$. Then the amplitudes of the elementary possibilities receive the values

$$\varphi_{\{H\}} = \frac{1}{\sqrt{3}},\ \varphi_{\{T_1\}} = \varphi_{\{T_2\}} = \frac{1}{2}\sqrt{\frac{2}{3}}, \tag{23}$$

yielding the required probabilities.

Both principles describe consistent mathematical conditions for probability amplitudes. But how can we compute substantial probability amplitudes? This is the content of the following two principles. From the second principle, we know that it is sufficient to calculate the amplitudes for all outcomes. From the first principle, it is clear to compute the amplitudes for the elementary possibilities only.

The **third principle** states that the amplitudes φ_F contribute equally in magnitude for all elementary possibilities, that is, the amplitudes are proportional to some constant times a complex number of magnitude one, namely

$$e^{\frac{i}{\hbar}S(F)} \text{ for all elementary possibilities } F \in \mathbf{F}. \tag{24}$$

The real-valued function $S(F)$ is called the *action* of the elementary possibility F.

This principle remembers at Laplace's *principle of indifference* where all outcomes should be equally likely assigned with unit one. Thus, the third principle can be viewed as a generalization that applies to elementary possibilities, and unit one is replaced by the set of complex numbers of magnitude one. If we define the phase as equal to zero, then we get back to Laplace's theory.

Originally, Feynman formulated this principle for probability amplitudes in the case of space-time paths: "The paths contribute equally in magnitude, but the phase of their contribution is the classical action (in units of $\hbar$); i.e., the time integral of the Lagrangian taken along the path[30]." Later he explained and summarized his rules as follows:

[30] Feynman [1948] p.9

> *The total amplitude can be written as the sum of amplitudes of each path - for each way of arrival. For every $x(t)$ that we could have - for every possible imaginary trajectory - we have to calculate an amplitude. Then we add them all together. What do we take for the amplitude for each path? Our action integral tells us what the amplitude for a single path ought to be. The amplitude is proportional to some constant times $exp(iS/\hbar)$, where S is the action for the path. If we represent the phase of the amplitude by a complex number, Planck's constant $\hbar$ has the same dimensions.*
> Feynman and Hibbs[31]

Thus Feynman's formulation for actions in phase space can be viewed as a particular case of our third principle. Please notice we make no further assumptions about the action except that it is real-valued. Hence, we are very flexible in describing physical problems outside space-time paths.

Our **fourth general principle** generalizes Feynman's principle for space-time routs[32] : "When a particle goes by some particular route, the amplitude for that route can be written as the product of the amplitude to go partway with the amplitude to go the rest of the way.“ This property goes back to Laplace, who investigated how to calculate the probability of events or experiments that can break down into a series of steps happening independently. Then the probability of the occurrence of all is the product of the probability of each.

Two possibilities F and G are called *independent* if their intersection is non-empty, and if the occurrence of one possibility does not affect the other one, that is, both have no influence on each other.

Mathematically, both possibilities are independent, if and only if their joint amplitude is equal to the product of their amplitudes:

$$\varphi_{F\cap G} = \varphi_F \, \varphi_G. \tag{25}$$

In our set-theoretic language of Section 4.1, parts of routs are subsets of the possibility space $\mathbf{F}$. The transition from a to b is the set $F = abC$, and the transition from b to c is the set $G = Abc$. If both transitions are independent of one another, then we multiply both corresponding amplitudes. We obtain $F \cap G = \{abc\}$ and $\varphi_{\{abc\}} = \varphi_F \, \varphi_G$.

In general, events are affected by previous events and thus are dependent. In accordance with Laplace experiments and classical probability theory, the probability amplitudes for independent possibilities are multiplied. In other words, the *multiply-and-add rule* carries over to complex numbers yielding the fundamental rule of quantum mechanics.

Notice that these principles would not be consistent if amplitudes would map into octonions or quaternions instead of complex numbers. These number systems are not associative and commutative.

In summary, these four principles serve to calculate the complex amplitudes for outcomes. They allow interference. Born's rule provides probabilities for all

[31] Feynman, Hibbs [1965], p.19
[32] Feynman Lectures [1963] p.3-4

outcomes. With Kolmogorov's axioms, we obtain classical probabilities for the non-elementary events. In most applications, the essential and challenging task is calculating the probability amplitudes. These are the relevant quantities, and sometimes it is not easy to calculate them.

Basic facts: The recipe for calculating probabilities:
Given an experimental set-up:

1. Define the possibility space **P** and the field **F**.
2. Define the sample space **O** of outcomes. They are subsets of the possibility space incorporating all internal possibilities.
3. Calculate the probability amplitudes for the possibilities 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 axioms the probabilities for the classical non-elementary events.

The possibility space **P** and the field of subsets **F** are defined similarly as in classical probability theory the sample space and the related field of subsets of the sample space. Moreover, the *multiply-and-add rule* holds for probability amplitudes as well. **The essential difference to other theories about probability is (i) that complex numbers are used from the very beginning, (ii) that possibilities and outcomes are different quantities, (iii) that internal possibilities, responsible for interference, are essential, and (iv) that we use the language of sets in contrast to many formulations of quantum theory**. Our theory can be viewed as a calculus with complex numbers that deliver numerical probabilities for outcomes based on experimental set-ups. 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 conflicting 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 our principles, see below. Experiments, classical or quantum ones, can be explained by using this recipe.

4.3 Reconstruction of Quantum Mechanics

We reconstruct Feynman's formulation[33] of quantum mechanics, which is known to be mathematically equivalent to Schrödinger's and Heisenberg's formulations. This reconstruction is rather simple since our probabilistic framework is closely related to Feynman's formulation. However, there are some differences: First of all, our language is set theory which significantly distinguishes between possibilities and outcomes. Secondly, classical probability and thermodynamics are incorporated. Thirdly, a massive object is not at several places simultaneously.

Let us assume that in the experimental set-up ABC the elementary possibilities of machine B are internal, such that the possibilities $ac := aBc$ are the outcomes. Moreover, let the possibilities $ab := abC$ and $bc := Abc$ be independent. 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 selected 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 (14) implies

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

Since $ab \cap bc = \{abc\}$, from (25) we get Feynman's[34] well-known formula (5):

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

The *superposition of probability amplitudes* (26) and (27) is the sum of several complex amplitudes, one for each route. This allows the cancellation of probability amplitudes, yielding the typical phenomena of *interference*. Both formulas provide the core of Feynman's theory, sometimes called *Feynman's sum-over-histories formulation*. 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 precisely 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.

Since all possibilities of B are internal and thus not given to the environment, the probability of detecting a particle in a and c must take account of all routes abc where b varies in B. Therefore, it is frequently stated that the quantum object seems to move on all possible routes simultaneously. In our approach the particle chooses only one route in the present, but with the tendency to move to states of higher probability.

In quantum mechanics, two fundamental concepts are striking. Firstly, the *superposition principle* which is discussed above. Secondly, *entanglement*

[33] Feynman [1948]

[34] See Section 2, Feynman [1948]

which is based on the quantum rule that composed systems are described by the tensor product space. The development of quantum informatics and quantum computing made it very clear that entangled multiple-particle states, which cannot be written as a product of single-particle states, are not exceptional but are the rule in quantum theory. Both concepts are discussed controversial depending on the used philosophic interpretation.

Entanglement is a structural and natural concept. There are easy understandable forms of entanglement, for instance, two welded coins or when a cat is entangled with a killing machine box. It is, however, far from being simple when two photons are entangled, but a large distance separates them. Einstein, believing in (3+1)-spacetime, referred to it with the phrase "spooky action at a distance". In a (3+3)-position-velocity space[35], however, the "spooky action at a distance“ of two entangled photons vanishes. The two photons are welded in the velocity space V, that is, they can be connected via a velocity $v \in V$. Notice that the notion of distance depends on the underlying space.

4.4 The Slit Experiment

Let us go through our probabilistic framework in terms of the double-slit experiment SWD, see Figure 2. The possibility space is

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

The internal possibilities are the slits a and b in W. Without detectors at the slits, it cannot be observed through which slit the particles goes in the present. Hence, the sample space of outcomes is the set

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

Now, we use the *multiply-and-add rule*. Our fourth principle implies

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

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

$$\varphi_{sWd_m} = \varphi_{sad_m} + \varphi_{sbd_m} \text{ for all } d_m \in D. \tag{31}$$

Inserting the concrete amplitudes for the possibilities using the third principle, we obtain the amplitudes for the outcomes. Born's rule provides the probabilities of the outcomes and using Kolmogorov's rules, we can calculate the probabilities for the non-elementary events. This shows the unbelievable simplicity of explaining the double-slit experiment within our framework of possibilities, internal possibilities, outcomes, and our categorization.

Now, we have a very general algorithm working with complex probability amplitudes assigned to possibilities which allow us to calculate probabilities for outcomes. This algorithm is applicable to

[35] Jansson [2017], see Section 4.13 for more details.

classical statistical problems as well as to problems with interference phenomena. This formalism is a key in our interpretation of the reconstruction of quantum mechanics, the latter known as the most fundamental physical theory. From our point of view, quantum theory is a timeless theory belonging to the categories structure and future. It uses the geometrical properties of experimental set-ups for calculating classical probabilities of outcomes. Quantum theory and classical probability theory are not different probability theories that are in contrast. In our approach, they complement one another. Dynamics happens in the present: There is no rest, and physical particles and systems have the tendency to move towards states of larger probability.

Feynman[36] wrote very honestly in his well-known book about quantum mechanics:

> *One might still like to ask: "How does it work? What is the machinery behind the law?" No one has found any machinery behind the law. No one can "explain" any more than we have just "explained". "No one will give you any deeper representation of the situation. We have no ideas about a more basic mechanism from which these results can be deduced. [...] Yes! Physics has given up. We do not know how to predict what would happen in a given circumstance, and we believe now that it is impossible - that the only thing that can be predicted is the probability of different events. It must be recognized that this is a retrenchment in our earlier ideal of understanding nature.* Feynman 1963

Feynman's point of view that "the only thing that can be predicted is the probability of different events" is supported much later by Fuchs and Peres[37]:

> *The thread common to all the nonstandard "interpretations" is the desire to create a new theory with features that correspond to some reality independent of our potential experiments. But, trying to fulfill a classical world view by encumbering quantum mechanics with hidden variables, multiple worlds, consistency rules, or spontaneous collapse without any improvement in its predictive power only gives the illusion of a better understanding. Contrary to those desires, quantum theory does not describe physical reality. It provides an algorithm for computing probabilities for the macroscopic events ("detector clicks") that are the consequences of our experimental interventions. This strict definition of the scope of quantum theory is the only interpretation ever needed, whether by experimenters or theorists.* Fuchs and Peres 2000

We have presented a short algorithm for computing classical and quantum probabilities, summarized above on one page. From a slightly different point of view, the main features are:

---

[36] Feynman Lectures [1963] p.1-16

[37] Fuchs, Peres [2000]

- We describe physics in terms of four categories: structure (background), future (prognostic view), present (momentary decisions, dynamics), and past (facts).
- We understand the basics of quantum mechanics as part of the structure and the future. *Born's rule* is a prognostic tool providing probabilities of future events[38].
- We distinguish between possibilities and outcomes. The presence of internal possibilities is responsible for interference.
- The amplitude of an elementary possibility is proportional to some constant times $exp(iS/\hbar)$, where S is the action. Both belong to the category structure.
- *Dynamics* is fundamentally stochastic and happens in the present such that (i) there is no rest, and (ii) there is the tendency to move toward states of higher probability.
- According to a *theorem of Hurwitz*, the field of complex numbers is the largest commutative field possessing indispensable properties of numbers. This is as a basic reason that quantum mechanics, the most fundamental physical theory, is a theory based on complex numbers[39] with a $U(1)$ symmetry.

In the recent publication of Wood[40], several historical remarks about the use, the necessity, and the evidence of complex numbers in physics are presented.

4.5 Some Philosohical Aspects

The double-slit experiment[41] with its diffraction pattern has been called "The most beautiful experiment in physics". The used experimental set-ups 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 experiment using large molecules called phthalocyanine. These molecules can be seen with a video camera exhibiting their macroscopic nature. The molecules are sent one at a time through the wall with slits, such that exactly one molecule only interacts with the set-up. At the screen of detectors, they arrive localized at small places. This behavior is typical for macroscopic objects, not for classical waves. Nobody has ever seen a collapsing wave. Moreover, the pictures of the molecules produced with a video camera demonstrate that the wave picture is dubious. Over a long period, the molecules, one after the other, build up into

[38] von Weizsäcker [2006]
[39] See Section 2.2 and Appendix A in Jansson [2017]
[40] Wood [2021]
[41] Crease [2002]

an interference pattern consisting of stripes. This distribution shows the same wave interference as if you drop two stones into a smooth pool simultaneously. Hence, it seems to be evidence that this big molecule might travel as a wave, in agreement with the widely celebrated *wave-particle duality*.

Strangely enough and frequently emphasized, the interference patterns in two-slit experiments disappear if we obtain information through which slit the molecule passes. Let a detector be positioned only at one slit. Then the interference pattern, where both slits are open without detectors, vanishes. Hence, the molecules passing through the slit without the detector seem to know that the other slit is equipped with a detector. This phenomenon is called *non-locality*: what happens in one location seems to affect what happens in a distant location instantaneously. Non-locality is a fundamental mystery of today's quantum mechanics.

There is another strange mystery called the *observer effect*, that is, observing effects *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". For example, the *von Neumann-Wigner interpretation*, also known under the name "consciousness causes collapse", consciousness is postulated to be necessary for the completion of quantum measurements.

Zeilinger, well-known for his pioneering experimental contributions to the foundations of quantum mechanics, gave an impressive talk in 2014 [42] "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 particles can go through two slits at the same time."

The basic postulates of quantum mechanics seem to be far away from sense experiences. Also Penrose supported this viewpoint. In his excellently written book[43] he writes on page 216 :

> *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*

[42] See for example YOUTUBE

[43] Penrose [2016, p.216]

numbers involving the "imaginary" square root of -1. Penrose 2016

Penrose gave, not unfounded, his famous book the title FASHION, FAITH and FANTASY.

These philosophical thoughts and insights are based on the foundation of standard physics using relativistic spacetime. Our categorization, the division into structure, future, present, and past, leads to a completely different way of looking at things. In the following, we want to examine this more closely.

Our conclusion: The slit experiment in 2012 with the large phthalocyanine molecules shows: (i) a molecule is not a wave, (ii) it supports 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.

Let us now show in detail how changing the experimental set-up changes the outcomes and the statistics. At first, we consider the experiment where slit b is closed. Then the possibility space is

$$SWD = \{sad_m : s \in S, a \in W, d_m \in D\}. \tag{32}$$

There are no internal possibilities. Therefore, the outcomes coincide with the elementary possibilities. The probability amplitude $\varphi_{sa} = 1$ since the other slit is closed. Hence,

$$\varphi_{sad_m} = \varphi_{sa}\varphi_{ad_m} = \alpha_m. \tag{33}$$

There is only one route. Born's rule implies $\Pr(sad_m) = |\alpha_m|^2$. Thus, we obtain a classical probability without any interference, as expected. Similarly, when slit b is closed, we obtain $\Pr(sbd_m) = |\beta_m|^2$ without any interference.

Now, we assume that both slits are open. Then the possibility space is defined in (28). The internal possibilities are the two slits a and b in the wall W. Hence, the sample space consists of the outcomes (29). 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 would pass with probability $\frac{1}{2}$ through each slit. Hence, we set $\varphi_{sa} = \varphi_{sb} = \frac{1}{\sqrt{2}}$. The probability amplitudes calculated by the multiply-and-add rule are

$$\begin{aligned} \varphi_{sWd_m} &= \varphi_{sa}\varphi_{ad_m} + \varphi_{sb}\varphi_{bd_m} \\ &= \tfrac{1}{\sqrt{2}}\alpha_m + \tfrac{1}{\sqrt{2}}\beta_m. \end{aligned} \tag{34}$$

Therefore, we get the probabilities

$$\begin{aligned} \Pr(sWd_m) &= |\tfrac{1}{\sqrt{2}}\alpha_m + \tfrac{1}{\sqrt{2}}\beta_m|^2 \\ &= \tfrac{1}{2}(\alpha_m + \beta_m)^*(\alpha_m + \beta_m) \\ &= \tfrac{1}{2}(\alpha_m^*\alpha_m + \alpha_m^*\beta_m + \beta_m^*\alpha_m + \beta_m^*\beta_m) \\ &= \tfrac{1}{2}\left(|\alpha_m|^2 + |\beta_m|^2\right) + \tfrac{1}{2}(\alpha_m^*\beta_m + \beta_m^*\alpha_m). \end{aligned} \tag{35}$$

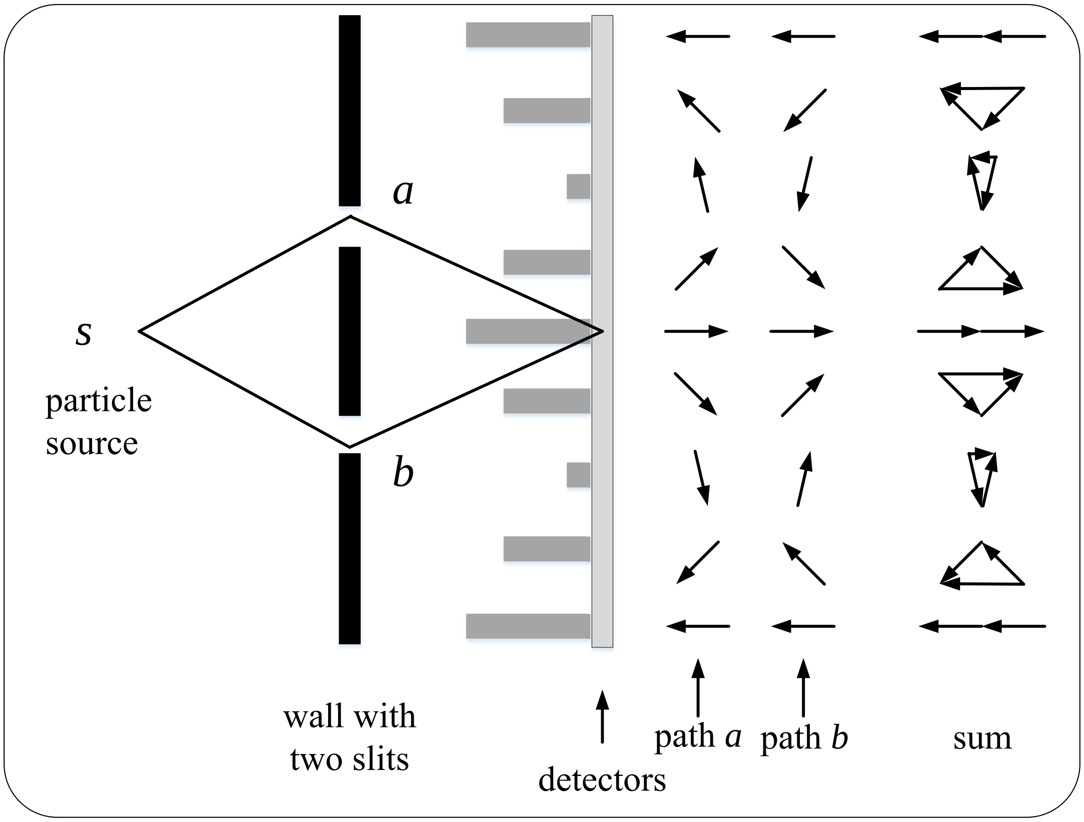


Figure 3: Schematic illustration of the double-slit experiment. The arrows represent the complex amplitudes for each path and their sum. Squaring the magnitude of the sum determines the corresponding probability. This leads to destructive and constructive interference, as displayed on the wall of detectors.

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 amplitudes with $\alpha_m = \beta_m$ we obtain from (35)

$$\Pr(sWd_m) = 2|\alpha_m|^2. \tag{36}$$

This doubles the classical probability, where only one slit is open. Hence, we have *constructive interference*. If $\alpha_m = -\beta_m$, the probability of finding the particle at detector d_m is

$$\Pr(sWd_m) = 0, \tag{37}$$

yielding *destructive interference*. For other combinations we obtain probabilities that are between both extreme cases.

Until now, we don't have the correct values for all amplitudes, such as α_m and β_m. We use the third principle. To calculate the amplitude for a particle with momentum p going from one position x_1 to another x_2, we need the classical physical action of this process. 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 $e^{ip(x_2-x_1)/\hbar}$, where $\hbar$ is Planck's constant.

We have described three different experiments that lead to different probabilities. In almost the entire literature, the question is asked how the particle knows which slits are open. The answer: magic, non-locality, wave-particle dualism, etc. The reason is the mental fixation on the particle. But it's the other way around. **The experimental set-up plays the primary role; the particle only plays a secondary role, which is limited to the interaction with the experimental set-up in the present. It tends to move to states of larger probability. Quantum mechanics is a theory that describes the structure of an experiment in its entirety with so-called amplitudes leading to probabilities via Born's rule**. In the following, this will become more clear.

Now, we want to discuss the case where we can get information about through which slit the particle passes. 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 in a future interaction arrives at detector d_m via slit b and detector d_a clicks, or both detectors d_a and d_b don't click.

The experimental set-up has changed. 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 4, 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

$$SWID = \{sad_ad_m, sbd_bd_m : \quad s \in S, a, b \in W, d_a, d_b \in I, d_m \in D\} \tag{38}$$

coincides with the sample space. 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.

The amplitude that a particle goes from source s via slit a to point m and detector d_a clicks is

$$\varphi_{sad_ad_m} = \varphi_{sa}\varphi_{ad_a}\varphi_{d_ad_m}. \tag{39}$$

For each outcome we have exactly one path. Our assumptions imply the probability amplitudes $\varphi_{ad_a} = 1$ and $\varphi_{sa} = \dfrac{1}{\sqrt{2}}$. Hence,

$$\varphi_{sad_ad_m} = \varphi_{sa}\varphi_{ad_a}\varphi_{d_ad_m} = \frac{1}{\sqrt{2}}\alpha_m, \tag{40}$$

which leads to the classical probability $\Pr(sad_ad_m) = 1/2|\alpha_m|^2$. Analogously, we obtain the classical probability $\Pr(sbd_bd_m) = 1/2|\beta_m|^2$. The fact that in

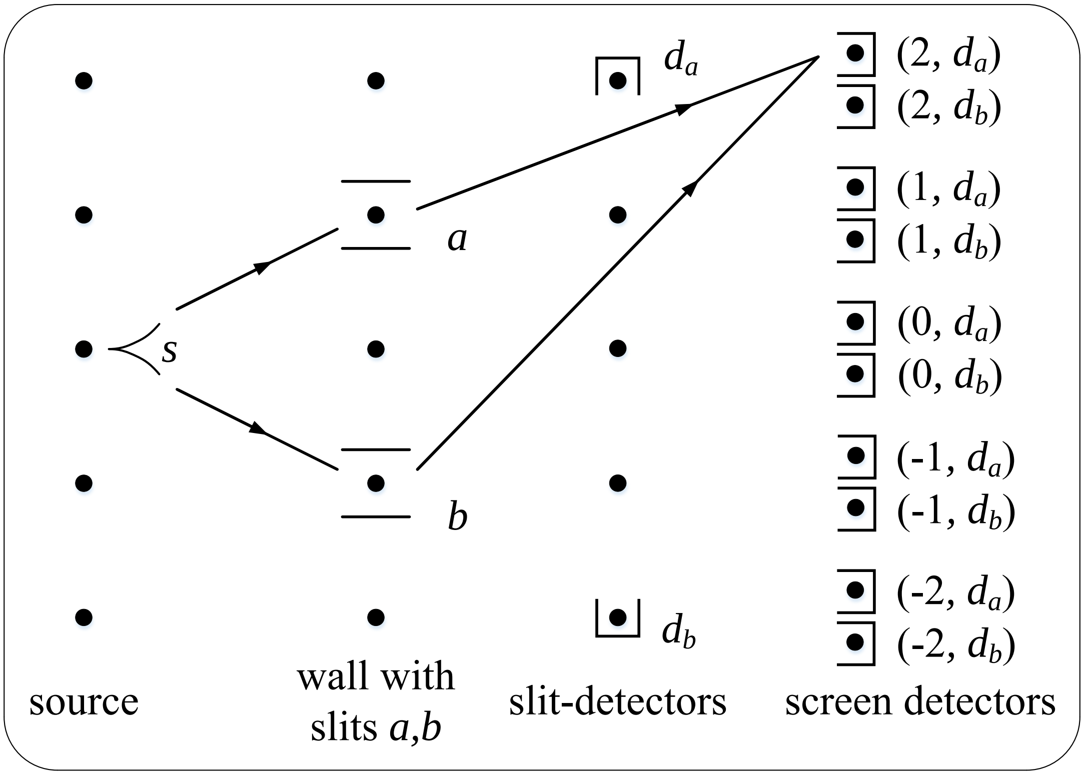


Figure 4: The double-slit experiment with slit-detectors. There are two paths for the event that a particle arrives at point 2 and detector d_a clicks. For the other points there are two paths as well.

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.

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.

It may also happen that a particle arrives at d_m via slit b and detector d_a clicks, or that a particle arrives at d_m via slit a and detector d_b clicks, or both detectors click or both don't click. These situations can be modeled as above and are left as an exercise. For example, if both detectors don't work, it is easy to show that we have interference as in the case without any detectors.

We want to explain how the strange philosophical aspects described above change when using our categorization.

- **Non-locality:** Whether there are detectors at the slits or not, or which slits are closed, are properties of the experimental set-up and belong to the category structure. Different set-ups imply different probability amplitudes. For example, we have seen that a detector at a slit gives further information outwards and changes the outcomes leading to other statistics. *Born's rule* provides probabilities that belong to the prognostic category future. The particle comes into play in the present and has no idea of the experimental set-up and the placed detectors. The only thing it does is to act according to the probabilities: There is no rest, and the particle tends to move towards states of larger probability. Hence, in our categorization, this weird non-locality does not appear. It is simply a structural property of the experimental set-up, not a strange behavior of the particle.

The next aspect mentioned above says that observing changes the reality. In other words, the act of observation may affect the properties of what is observed.

- **Observing:** In our framework, observing is described in terms of detectors belonging to the experimental set-up. The addition of further detectors, their functioning and their reliability is responsible for the occurrence of interference, as we have seen. It should be realized that an observer, who takes note of what happens, is entirely unnecessary. Nothing weird happens. The set-up of the experiment decides which of the mathematical models described above applies to the double-slit experiment.

The third aspect is the widely celebrated *wave-particle duality*, which is deeply embedded into the basics of quantum theory.

- **Wave-particle duality:** It states that every particle may be described as either a particle or a wave. The complete information about a particle is encoded in its wave function, which evolves according to the Schrödinger equation. Wave-particle duality expresses the incapability to describe the behavior of quantum objects with the standard physical concepts "wave" and "particle". **In our framework, we can: the particle concept continues, and the wave is not like a water or a sound wave. It is simply a probability distribution according to the experimental set-up**.

  We mention a further series difficulty of the wave-particle 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. Moreover, in quantum mechanics, two-state systems are frequently discussed. These are systems that can exist in a superposition of two mutually exclusive 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 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 set-up. These machines are globally positioned in a large area. Sometimes this give the impression that quantum mechanics is non-local as described above.

The wave concept is based on the superposition principle. Perhaps, the most fundamental question in quantum mechanics is: "What really happens in a superposition when only one particle is the experiment¿' This question is usually answered with weird statements such as: "The particle is in several states at once" and " The particle interferes with itself“. This is a view with a focus on the tiny particles. As frequently mentioned, our focus is on probabilities derived from the experimental set-up where active particles are irrelevant since they interact with the set-up only in the present.

- **Superposition:** Usually, the superposition principle means that states are described by vectors in a Hilbert space, and that each linear combination of states is a state again. In other words, every vector of a Hilbert space corresponds to a state. In our probabilistic approach we have not defined states. We work with set theoretical concepts possibilities and outcomes which live in the category structure. There, *superposition* means that the possibilities of one machine can be expressed in terms of the possibilities of other machines via probability amplitudes. For example, the possibility that a particle will hit a detector in the wall of the detectors in the future can be expressed by all possible paths through the wall of the slits to this detector as the end point via probability amplitudes. A particle would be able to choose this path *as well as* another path. Only in the present it has to choose *either this or that* path.

What about Zeilinger's mysteries?

- **Zeilinger's mysteries:** How do we resolve Zeilinger's quantum mysteries? Our categorization and explanation of the double-slit experiment below break down the illusion that a particle is a wave and can be at several places simultaneously. This supports experimental observations: an atom being at different places simultaneously has never been 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 that carefully distinguishes between outcomes and possibilities. The experimental set-up, 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. The molecule's property is the local interaction with the machines in the present. The patterns of the double-slit experiment become facts of the past. There is no mystery. Mystery occurs because most well-known arguments are based on pushing the "local“ properties of molecules in the foreground, and not the global aspects of experimental set-ups.

4.6 Dice Unlike Any Dice

The physicist Anthony Zee[44], well-known for his publications in quantum field theory, particle physics, and other topics in theoretical physics, has used the

[44]Zee [2015], p.141

title of this section in his book "Fearful Symmetry". He started this section with the words

> *Welcome to the strange world of the quantum, where one cannot determine how a particle gets from here to here. [...] When a die is thrown, the probability of getting a 1 is 1/6. The probability of getting a 2 is, of course, also 1/6. Now, consider the following question: What is the probability of getting a 1 or a 2 in one throw? The answer is evident to gamblers and non-gamblers alike: The probability is 1/6 + 1/6 =1/3. In everyday life, to obtain the probability of either A or B occurring, we simply add the probability of A occurring and the probability of B occurring.*
>
> *The quantum die is astonishingly different. Suppose we are told that for the quantum die the probability of throwing a 1 is 1/6, and the probability of throwing a 2 is also 1/6. In contrast to what our experience with ordinary dice might suggest, we cannot conclude that the probability of getting either a 1 or a 2 in one throw is 1/3! It turns out that the probability of throwing a 1 or a 2 can range between 1/3 and 0!*

It seems that quantum theory gives other results than the classical probability theory. Zee is an expert in quantum theory, and what he writes is widely accepted and is common sense. Apparently, Zee views classical probability theory as incompatible with quantum theory, that is, both theories are entirely different and handle different statistical applications and problems. This understanding is contrary to our unified probabilistic recipe.

Let us apply our recipe to dies. What is the probability of getting a 1 or a 2 in one throw? A die has 6 elementary possibilities $1, 2, 3, 4, 5, 6$. For each possibility, we set the action equal to zero and choose the probability amplitude equal to $1/\sqrt{6}$. There are no internal possibilities. The outcomes coincide with the elementary possibilities, yielding with Born's rule the probabilities $1/6$ for each outcome. Then, using the classical Kolmogorov rules, we get the probability $1/3$ for obtaining a 1 or a 2 in one throw? There is nothing strange.

However, there is a fundamental difference between standard quantum theory and our probabilistic framework, although we have reconstructed quantum theory. The significant difference is the clear distinction between internal possibilities, possibilities, and outcomes. In Section 5, we show how our recipe reconstructs the fundamental theory of statistical thermodynamics. There, microstates correspond to elementary possibilities, and macrostates correspond to outcomes. Similarly, our recipe works satisfactorily and reasonably for various classical applications.

4.7 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 occur in different orders simultaneously. In recent literature[45], it

[45] See for example Goswami et al. [2018], Wolchover [2021]

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. There, 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 simultaneously, both causal cases exist at the same time, thus destroying causality.

Experimentally, the quantum switch can be realized as an optical set-up 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 apply 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

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle_c + \frac{1}{\sqrt{2}}|1\rangle_c, \tag{41}$$

then the output state of the system is in the superposition

$$|\Psi\rangle = \frac{1}{\sqrt{2}}BA|\phi\rangle_t \otimes |0\rangle_c + \frac{1}{\sqrt{2}}AB|\phi\rangle_t \otimes |1\rangle_c, \tag{42}$$

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

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.

Not surprisingly, in our approach causality is not violated. Why? 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 not violated in our categorization.

4.8 Our World on Three Pages

Most textbooks in physics are filled with equations of motion described as differences of some physical quantities. A completely different approach in physics is the *action principle*. From the mathematical point of view, the most simple form of an action is defined as a real-valued function that has trajectories, also called paths or histories, as its arguments. If a particle moves in spacetime, the action is calculated as follows: We subtract the potential energy from the kinetic energy, and then we sum up (integrate) these energy differences over the time interval corresponding to the path. Time t can be viewed merely as a geometrical parametrization of the path, not as our perception of a physical flow corresponding to reality.

The action allows to derive the equations of motion. In classical mechanics, the path followed by a system is that one that makes the action stationary. The symmetry of spacetime demands that the equations of motion must hold in each reference frame. They must be *covariant* with the *Lorentz transformation*, that is, when applying a Lorentz transformation, the quantities on both sides of the equation change, but such that both sides stay equal. In contrast, *invariant* quantities do not change when applying these transformations. The equations of motion are covariant, but the action is invariant with respect to Lorentz transformations[46]. Hence, the actions of the elementary possibilities in the **third principle** (24) are invariant for these spacetime transformations and thus fit into our timeless probability recipe.

The action is an additive quantity, and as soon as we can describe a new area in physics in the form of an action, this action is added to the whole action expression. Then we get a single formula, called the *path integral*, that could be similar to the expression

$$S = \int dx \sqrt{g} \left[\frac{1}{G} + \frac{1}{g^2} F^2 + \overline{\psi} \widehat{D} \psi + (D\psi)^2 + V(\psi) + \overline{\psi} \phi \psi \right]. \tag{43}$$

Unfortunately, understanding this formula requires years of intensive study of physics[47].

It has been found that the principle of action is universal and can be applied to all physics. In other words, the entire physical world is based on a fundamental quantity, the action. This affects classical mechanics, Maxwell's equations, the ten equations of general relativity, and quantum mechanics. The important symmetries in physics (spacetime translation, gauge symmetry, etc.) are symmetries of the action. The continuous symmetries of the action imply conservation laws.

It seems to be very natural to believe that the the single purpose of classical time-symmetric theories, like mechanics in spacetime, electromagnetism, or gravitation, is to compute an action functional, which produces the phases of probability amplitudes. The probability amplitudes form the basis of statistical motion, as described in our probabilistic algorithm. There is a close relationship to the *Wiener integral* for solving problems in diffusion and to *Brownian motion* yielding non-smooth zigzag paths. In fact, the Feynman principles are referred frequently to the work of Norbert Wiener on Brownian motion in the early 1920s. Since Feynman sums up all paths, his approach

[46]Zee [1993], Chapter 7, Zwiebach [2004], Chapter 5

[47]Zee [1993], The heading of this section is a variation of the german title in Zee's book "Die ganze Welt auf einer Serviette", page 134.

is well-known under the name *Feynman path integral*. 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.*[48]

There is a fascinating, beautifully written, and comprehensive physic book by Schiller [49] where he investigated in part IV "The Quantum of Change". This book is highly recommendable for students in engineering. His starting thesis is:

- the action values S_1 and S_2 between two successive events of a quantum system cannot vanish. They satisfy the inequality $|S_2 - S_1| \geq \frac{\hbar}{2}$.

This *minimum action principle* is in complete contrast to classical physics, but has never failed a single test, as pointed out in his book. Based on the quantum of change, Schiller deduced several consequences that cannot be found in other textbooks but agree with Feynman's view:

- In nature, there is no *rest*.
- In nature, there is no perfectly straight or perfectly *uniform motion*.
- Perfect *clocks* do not exist.
- Motion *backward in time* is possible over microscopic times and distances.
- The *vacuum* is not empty.
- Photons have no position and cannot be localized.
- Microscopic systems behave randomly.
- Light can move faster than the speed of light c.

Now, unbelievable many applications of path integrals in physics are known, including the harmonic oscillator, particles in curved space, Bose-Einstein condensation and degenerate Fermi gases, atoms in strong magnetic fields and the polaron problem, quantum field-theoretic definition of path integrals, or string interactions. The Feynman path integral is known as a candidate theory for the quantum gravity problem. In the context of quantum cosmology, some investigations about the start of our universe using path integrals are known. One can find many details in the comprehensive book of Kleinert[50].

[48] `https://en.wikiquote.org/wiki/Freeman_Dyson`
[49] Schiller [2016]
[50] Kleinert [2009]

In summary, the fundamental quantity *action*, appropriately applied to our probability recipe in Section 4, provides an algorithm that allows us to describe almost all experimental results but avoids well-known paradoxes. This algorithm might be viewed as a program for solving experimental problems, not for explaining our world ontologically.

5 Reconstruction of Thermodynamics

> *A theory is the more impressive the greater the simplicity of its premises is, the more different kinds of things it relates, and the more extended is its area of applicability. Therefore the deep impression which classical thermodynamics made upon me. It is the only physical theory of universal content concerning which I am convinced that within the framework of the applicability of its basic concepts, it will never be overthrown.* Albert Einstein, Autobiographical Notes (1946)

It is an important touchstone for our probability theory to reconstruct thermodynamics, this physical theory of universal content. It turns out that thermodynamics can be viewed as a straightforward application of our probabilistic recipe described in Section 4.

Statistical thermodynamics, a large area of statistical mechanics, was developed primarily by the Austrian physicist Boltzmann (1844 - 1906), but also other scientists contributed to it, among them Maxwell and Gibbs. Boltzmann applied statistical methods to the controversial discussed atomic hypothesis.

Usually, in statistical mechanics, huge numbers of constituents are considered. For a system composed of a large number of particles, say of the order of Avogadro's number $\approx 10^{23}$ which corresponds to 1 mole of molecules, it is not possible to follow their trajectories, regardless of whether they even exist. Notice that the number of all grains of sand on all beaches in our world is about 10^{19}. Moreover, we have no accurate initial conditions, namely exact positions and momenta of each particle, required for such calculations. Therefore, one can deduce only statistical descriptions in thermodynamics. The question arises whether there exist few statistical macroscopic parameters that determine approximately the thermodynamical system. Actually, such parameters exist.

Thermodynamics, however, with its basic concepts like the second law, entropy, and the "Time's Arrow", is discussed controversially; see for example Ben-Naim[51]. He writes that "Time does not feature in thermodynamics in general, nor in entropy", and "Reading through the entire book by Eddington[52], you will not find a single correct statement on the thermodynamic entropy". We would like to show that our approach leads to a new understanding of thermodynamics.

[51] Ben-Naim [2018] Chapter 1
[52] Eddington [1927]

5.1 Two-State Systems

It is beneficial, but not necessary, if the reader has some knowledge of statistical thermodynamics. There are plenty of books and articles on this subject. For a nice introduction, we mention Penrose[53], and moreover four textbooks[54] which are well-suited for engineers.

We start this section with an introductory example. Suppose that n indistinguishable molecules are placed in a box consisting of N cells. There are

$$\Omega(n, N) = \binom{N}{n} = \frac{N!}{n!\,(N-n)!}. \tag{44}$$

possibilities of dividing n objects in N cells: N cells where the first molecule can be placed, $N-1$ cells where the second one can be placed, and finally $N-n$ cells where the last one can be positioned. Hence, we get $\frac{N!}{(N-n)!}$ possibilities. But the constituents are indistinguishable such that the $n!$ configurations are not distinct, leading to the remaining denominator. The number $\Omega(n, N)$ is called the *multiplicity* of the *macrostate* (n, N). Its distinct configurations are called the accessible *microstates*.

Instead of always working with large multiplicities, it is convenient to define the entropy

$$S_B(n, N) = k \ln \Omega(n, N). \tag{45}$$

This form is called *Boltzmann entropy*. It is placed on his gravestone. The logarithm makes large numbers manageable, but more important is that the entropy of two independent systems must be added to get the entropy of the combined total system. Entropy is an additive quantity like energy.

Suppose we have a second box consisting of M cells that contains m molecules. Then the number of configurations, when both boxes are kept separate, is the product

$$\Omega(n, m, N, M) = \Omega(n, N)\,\Omega(m, M), \tag{46}$$

since for each microstate of the first system there are $\Omega(m, M)$ microstates of the second system. Hence, the entropy is in this case

$$S_B(n, m, N, M) = k \ln \Omega(n, m, N, M) = S_B(n, N) + S_B(m, M). \tag{47}$$

The constant factor k is purely conventional and is chosen according to the application. Boltzmann has chosen a special constant k_B for k that relates the average relative kinetic energy of gas molecules with the temperature. The multiplicity, and equivalently the entropy, express in some sense the *disorder* or *uncertainty* of the system in a given macrostate. If the system has a small multiplicity, then the system can be only in a few microstates and has low entropy.

[53] Penrose [2005], Chapter 27

[54] Schroeder [1999], Gould, Tobochnik [2010], Schwarz [2017], Swendsen [2020]

macrostates	microstates
$(3,0)$	$(p_1p_2p_3)$
$(2,1)$	$(p_1p_2q_1)(p_1p_3q_1)(p_2p_3q_1)(p_1p_2q_2)(p_1p_3q_2)(p_2p_3q_2)$
$(1,2)$	$(p_1q_1q_2)(p_2q_1q_2)(p_3q_1q_2)$

Table 1: Two boxes with 3 macrostates and 10 microstates. The number of accessible microstates, the multiplicity, varies between one and six.

What happens if it is allowed that the molecules can be interchanged between both boxes? What are the microstates of both boxes in this case? Let us consider the case $n = 2, N = 3, m = 1, M = 2$. There are three macrostates of the combined boxes. The first one contains all three molecules in the first box, denoted by $(3, 0)$, the second one is $(2, 1)$, and the third one is $(1, 2)$. Other combinations are not possible. Each macrostate consists of a set of configurations, that is, of accessible microstates. For example, $(p_1p_3q_2)$ denotes the state where one molecule occupies the first cell p_1 in the first box, a second molecule occupies the third cell p_3 in the first box, and the third particle is in the second cell q_2 of the second box. All macrostates and microstates are displayed in Table 1.

There may be many accessible microstates for each macrostate. The *total multiplicity* Ω_{tot} is the sum over the multiplicities of all macrostates, that is, the sum over all microstates. In our example, we have 10 microstates yielding the total multiplicity 10. The *total entropy* of the system is $S_B = k \ln \Omega_{tot}$.

A macrostate is observable. We have to separate both boxes, thus disallowing the interchange of molecules. Separating both boxes acts as a constraint. For example, if we separate both boxes and observe the macrostate $(2, 1)$, then the first box has only one macrostate (2) containing three microstates $(p_1p_2), (p_1p_3), (p_2p_3)$, and the second box has only one macrostate (1) containing two microstates $(q_1), (q_2)$. Hence, **removing any constraint of an isolated combined system will increase multiplicity and entropy**. This is one formulation, perhaps not the most known version, of the *second law of thermodynamics*[55].

The thermodynamic *equilibrium*[56] is the macrostate with the greatest multiplicity, or equivalently with the highest entropy. In our example, it is the state $(2, 1)$. If $N = M$ and $n = m$, then it is easy to show that the equilibrium is the macrostate (n, n, N, N) where both boxes have an equal number of molecules. If the number of cells and the number of molecules are different, then the boxes have equal concentrations $\widetilde{n}/N$ and $\widetilde{m}/M$ in the equilibrium state. This property justifies the name equilibrium.

So far, all these thermodynamic quantities, namely microstates, macrostates, multiplicity, entropy, equilibrium, the second law of thermodynamics, and disorder belong to the category structure. They describe the deterministic structure of the experimental set-up without any dynamics or uncertainty.

The fundamental principle in statistical thermodynamics states that **all microstates of a system are equally probable**. This is a probabilistic

[55] Ben-Naim [2018] Chapter 4
[56] Attard [2002], Chapter 1

statement and thus belongs to the category future. It follows that the probability of a macrostate is the multiplicity of this macrostate divided by the total multiplicity Ω_{tot}.

But what about the dynamics of a thermodynamic system? We claim that **a system is never at rest and has the tendency to move from microstate to microstate toward macrostates of larger probability**. Macrostates that are supposed to occur more frequently occur more frequently. However, this is a tendency, and the system may also move to a macrostate with low probability. Even if the macrostate is the equilibrium, the system may move to other macrostates with lower probabilities. We speak of *fluctuations*. These weak statements about motion belong to the category present.

In our example, we have 10 microstates, and each has the probability 1/10. The microstates are mutually exclusive. Thus the macrostate $(3, 0)$ has probability 1/10, the macrostate $(2, 1)$ has the probability 6/10, and $(1, 2)$ has the probability 3/10. Hence, on the average we expect to observe the system in state $(3, 0)$, $(2, 1)$, or $(1, 2)$ in about 1/10, 6/10, or 3/10 of the observations, respectively. It is simple to write a program that simulates transitions between the microstates in agreement with these probabilities. Obviously, a system starting in a low probability macrostate and moving to an equilibrium, sometimes returns to this initial state. Hence, thermodynamics is not irreversible. In particular, the entropy may decrease. Depending on the experiment, a return to the initial state is not impossible but may be extremely improbable. The widely celebrated association of entropy and the idea of the "times arrow" is barely comprehensible.

To illustrate the basics above, we could also examine each other two-state system, such as a paramagnet, spin, polarization, coin toss, or mixing colors. Two-state systems are considered universal by some scientists. Very early in the fifties, Weizsäcker[57] formulated two principles of his *ur theory*:

- **Principle of alternatives:** *Physics reduces to measurement outcomes, the only available quantities. Thus physics is best formulated based on empirical decidable alternatives. Alternatives describe mutually exclusive states, events, outcomes, possibilities, or facts. They either happen or do not happen, but two or more alternatives cannot happen simultaneously.*

Then he restricted physics further, stating his

- **Ur hypothesis**: *All alternatives can be constructed from binary alternatives.*

Hence, physics can be defined entirely in terms of binary alternatives and their symmetries[58]. In other words, Weizsäcker developed the view that "all physical models can completely be derived from the information contained in an ur". For more details, see Section 6.

This small section already provides the basic machinery of statistical thermodynamics. In the following sections, we will deal more generally with ther-

[57] von Weizsäcker [1955], von Weizsäcker [1958]

[58] See also von Weizsäcker [1988], von Weizsäcker [1992], von Weizsäcker [2006]

modynamics, its reconstruction from the principles in Section 4.2, and its generalizations up to optics in Section 5.5.

5.2 Reconstruction

Each thermodynamic system is fundamentally described by a set of *microstates*, which are specified by the states of each constituent of the system, such as positions and momenta or the quanta of energy for each constituent. In other words, it is a specific microscopic configuration of a system where all possible microscopic variables are fixed. The microstates form distinguishable alternatives, that is, they either happen or do not happen in the present, but two or more of the microstates cannot occur simultaneously.

Macrostates refer to the state of the system as a whole. There are only a few macroscopic variables such as the total energy E, pressure P, volume V, temperature T, the total number N of gas molecules, or magnetization M. Many experiments can be described by the three macroscopic variables E, V, N only. In contrast to few macrostates, the number of microstates may be huge, and they contain all details of the system. A macrostate emerges by fixing the value of every macroscopic variable. The macroscopic variables can be observed or measured in contrast to the large number of microstates. Macrostates form a partitioning of the set of microstates.

The obvious way to connect the concepts of thermodynamics to our recipe for calculating probabilities is to identify the microstates as elementary possibilities. The macrostates, as observable states, correspond to the outcomes. These are characterized by their accessible microstates. They define the elementary events of the corresponding classical sample space. The accessible microstates of a macrostate are indistinguishable with regard to macroscopic observations, but microstates belonging to different macrostates are macroscopically distinguishable. In the following, we will use the equivalent terms microstates and elementary possibilities as well as macrostates and outcomes optionally. Moreover, we write shortly $\mathbf{M}$ for a macrostate and μ for a microstate.

Next, we shall reconstruct the probabilities for macrostates with our probabilistic framework described in Section 4.2. The **third principle** states that all elementary possibilities contribute equally in magnitude, that is, the amplitudes are proportional to some constant times a complex number of magnitude one. Given a macrostate $\mathbf{M}$, then its accessible microstates μ have the amplitudes

$$\varphi_\mu = \text{const } e^{\frac{i}{\hbar}S(\mu)} \text{ for all microstates } \mu \in \mathbf{M}. \tag{48}$$

Since we have no further knowledge about actions of the constituents, it seems to be natural to set the action $S(\mu) = 0$ for all microstates. The action belongs to the category structure. Moreover, we set

$$\text{const} = \frac{1}{\sqrt{\Omega_{tot}}\sqrt{\Omega(\mathbf{M})}}. \tag{49}$$

Then

$$\varphi_\mu = \frac{1}{\sqrt{\Omega_{tot}}\sqrt{\Omega(\mathbf{M})}}\, 1. \tag{50}$$

The microstates are pairwise disjoint, and using the second principle (superposition) we get the probability amplitude of the macrostate $\mathbf{M}$:

$$\varphi_{\mathbf{M}} = \sum_{\mu \in \mathbf{M}} \varphi_\mu = \Omega(\mathbf{M}) \frac{1}{\sqrt{\Omega_{tot}}\sqrt{\Omega(\mathbf{M})}} = \sqrt{\frac{\Omega(\mathbf{M})}{\Omega_{tot}}}. \tag{51}$$

With Born's rule we obtain by computing the square of the magnitude of probability amplitudes the classical probabilities for the outcomes, namely the multiplicity of this macrostate divided by the total multiplicity Ω_{tot}. These probabilities belong to the prognostic category future.

Notice, in our derivation, we did not apply the probabilistic principle of indifference belonging to the category future. Instead, we set the action of all elementary possibilities equal to zero, which is a statement about the experimental set-up.

5.3 Entropy

Entropy is a fundamental physical quantity in thermodynamics, statistical mechanics, quantum theory, and information theory. Unfortunately, there is a problem because of the various definitions of entropy. Two interesting publications about different forms of entropy are given by Schwartz[59] and Županovi and Kui[60]. In this and the next section, we try to establish some relationships between these definitions.

The *Boltzmann entropy* is defined for macrostates $\mathbf{M}$ as

$$S_B(\mathbf{M}) = k \ln \Omega(\mathbf{M}), \tag{52}$$

and the *total entropy* of the system is

$$S_B = k \ln(\Omega_{tot}), \quad \Omega_{tot} = \sum_{\mathbf{M}} \Omega(\mathbf{M}). \tag{53}$$

The entropy belongs to the category structure and can be rewritten as:

$$S_B = k \frac{1}{\Omega_{tot}} \sum_{\mathbf{M}} \Omega(\mathbf{M}) \ln(\Omega_{tot}). \tag{54}$$

Since all microstates of a system are equally probable, we can write

$$S_B = k \sum_{\mathbf{M}} \Pr(\mathbf{M}) \ln(\Omega_{tot}). \tag{55}$$

[59]Schwartz [2019], Chapter 6
[60]Zupanovi, Kui [2018]

Since

$$\ln(\Omega_{tot}) = \ln(\Omega(\mathbf{M})) + \ln\frac{\Omega_{tot}}{\Omega(\mathbf{M})} = \ln(\Omega(\mathbf{M})) - \ln(\Pr(\mathbf{M})), \tag{56}$$

we finally obtain

$$S_B = \sum_{\mathbf{M}} \Pr(\mathbf{M}) S_B(\mathbf{M}) - k \sum_{\mathbf{M}} \Pr(\mathbf{M}) \ln \Pr(\mathbf{M}). \tag{57}$$

Notice that we have two mathematical identities for the entropy S_B. However, these are not contentwise identities since probabilities are prognostic statements, whereas multiplicities and the Boltzmann entropy are structural properties of the experiment. The first term in this sum is the mean value of the Boltzmann entropy. The second term applies to macrostate probabilities only. It describes the uncertainty due to the various macrostates of the system. If this distribution is sharply peaked around the equilibrium state, then its probability is almost one. Since $\ln 1 = 0$, the second term and the uncertainty vanish. Moreover, the total entropy is nearly equal to the mean value of the Boltzmann entropy.

Frequently in the literature, the macrostate probability is expressed as an exponential of the entropy:

$$\Pr(\mathbf{M}) = \frac{\Omega(\mathbf{M})}{\Omega_{tot}} = \frac{1}{Z} e^{S_B(\mathbf{M})/k}, \text{ where } Z = \Omega_{tot} = \sum_{\mathbf{M}'} e^{S_B(\mathbf{M}')/k}. \tag{58}$$

Formally, the second term in (57) applies to any probability distribution. It was originally presented by Gibbs. Later, Shannon derived this expression in his theory of communication and information.

The second law of thermodynamics is frequently formulated in terms of entropy. There are, however, several different opinions about the second law. For example, the well-known free encyclopedia Wikipedia[61] introduces the second law with the words:

> *The second law of thermodynamics establishes the concept of entropy as a physical property of a thermodynamic system. Entropy predicts the direction of spontaneous processes and determines whether they are irreversible or impossible, despite obeying the requirement of conservation of energy, which is established in the first law of thermodynamics. The second law may be formulated by the observation that the entropy of isolated systems left to spontaneous evolution cannot decrease, as they always arrive at a state of thermodynamic equilibrium, where the entropy is highest. If all processes in the system are reversible, the entropy is constant.*

In our approach, this fundamental law has an entirely different meaning. It belongs to the category structure and says that when we remove in the experimental set-up of an isolated system any constraint, then multiplicity and

[61] $[https://en.wikipedia.org/wiki/Second_law_of_thermodynamics]$

entropy will increase. Contrary, if we add new constraints, then the entropy will decrease. We have seen in Section 5.1 that a system starting in a low probability macrostate and moving to the equilibrium can sometimes return to the initial state. This is not impossible but highly improbable. It is just as unlikely as an archaeopteryx will grow out of the earth. But according to the theory presented here, that would be possible. Thermodynamics is not irreversible. In particular, the entropy may decrease.

5.4 Quantum Entropy

In quantum theory a fundamental postulate says that quantum states are state vectors ψ_j in a complex Hilbert space. Frequently, so-called *mixed states* are considered, where only classical probabilities $P_j = \Pr(\psi_j)$ for some orthogonal state vectors are known. These mixed states are described by the density matrix

$$\rho = \sum_j P_j \psi_j \psi_j^T. \tag{59}$$

The *von Neumann entropy* is defined as

$$S_N = -k \ \mathrm{Trace}(\rho \ln \rho). \tag{60}$$

The density matrix is a positive semi-definite, Hermitian operator of trace one. It is not hard to prove that S_N is well-defined, basis independent, and can be written in the form

$$S_N = -k \ \sum_j \langle \psi_j | \rho \ln \rho | \psi_j \rangle = \sum_j P_j \ln P_j. \tag{61}$$

This is the same mathematical form as the right-hand side in formula (57), and formally equivalent with the Gibbs and the Shannon entropy, as already mentioned. For a pure state, such that $P_j = 1$ for some j and zero otherwise, the von Neumann entropy is zero. In a pure state, we have certainty. Otherwise, it is a basis-independent measure of uncertainty.

The von Neumann entropy is the standard in quantum information theory. Our probabilistic framework allows a completely other way to look at entropy from the point of view of quantum theory. The *Boltzmann entropy* (45) can be viewed as a scaled multiplicity. The multiplicity of a macrostate $\mathbf{M}$ is the number of its accessible microstates $\mu \in \mathbf{M}$, and thus can be written as

$$\Omega(\mathbf{M}) = \sum_{\mu \in \mathbf{M}} 1. \tag{62}$$

Our **third principle** replaces the *principle of indifference* by substituting the 1 with a complex number of magnitude one. If we apply this rule to the multiplicity, we get a *complex multiplicity* for a macrostate $\mathbf{M}$:

$$\Omega_{\mathbb{C}}(\mathbf{M}) = \sum_{\mu \in \mathbf{M}} e^{\frac{i}{\hbar} S(\mu)}, \tag{63}$$

where $S(\mu)$ denotes the action of microstate μ. The action allows to incorporate geometrical properties of the experimental set-up. When we have no additional geometrical information as in statistical thermodynamics, the action is set equal to zero, and we get back classical multiplicities.

In the case of a slit experiment, the paths correspond to the microstates. The outcomes, or equivalently the macrostates, correspond to the set of paths from the source to a particular detector. Then, except for one real constant, the complex multiplicity coincides with Feynman's path integral. This is a surprising fact. In other words, we can identify the complex multiplicity with Feynman's path integral. Conversely, the usual multiplicity is a special case of the path integral.

If we replace in Boltzmann's entropy the multiplicity by the complex multiplicity, it follows that

$$S_{B,\mathbb{C}}(\mathbf{M}) = k \ \ln \ \Omega_{\mathbb{C}}(\mathbf{M}) = k \ \ln \sum_{\mu \in \mathbf{M}} e^{\frac{i}{\hbar} S(\mu)}. \tag{64}$$

We call it the *quantum Boltzmann entropy.*

The complex multiplicity is a complex number, and can be written in the form

$$\Omega_{\mathbb{C}}(\mathbf{M}) = r \ e^{\frac{i}{\hbar} S}, \tag{65}$$

where r and S are real numbers. Thus, the quantum Boltzmann entropy takes the form

$$S_{B,\mathbb{C}}(\mathbf{M}) = k \ \ln \ r + \ k \frac{i}{\hbar} S. \tag{66}$$

In the extreme case where the action $S(\mu) = 0$ for all accessible microstates $\mu \in \mathbf{M}$, it follows that $S = 0$, and $S_{B,\mathbb{C}}(\mathbf{M})$ coincides with the Boltzmann entropy. It is natural to interpret S as an average action of a macrostate $\mathbf{M}$ and r as an average real multiplicity. In other words, this quantum entropy gives back an average Boltzman entropy $k \ \ln \ r$ on the real axis and an average action S on the imaginary axis.

5.5 Light Reflection

In this section, we answer the question stated in the preface: What has statistical thermodynamics and the reflection of light on a mirror have in common?

The seemingly simple problem of how light is reflected by a mirror is usually solved with the well-known *ray model of light* in optics, see Figure 5. It says that the mirror reflects light in a way such that the angle of incidence is equal to the angle of reflection. Moreover, the length of the mirror, as well as the right and the left end of the mirror, do not influence the light that reaches the detector. This model describes light in terms of rays and holds in many practical situations.

The experimental set-up is as follows: at a source, the light of one color is emitted, and at another point, there is a photomultiplier for detecting light.

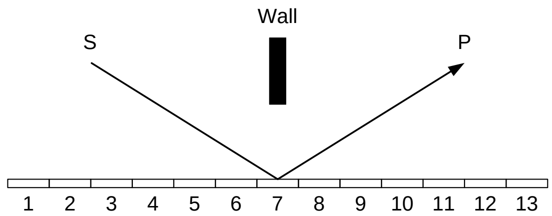


Figure 5: The classical view of the ray model: the mirror reflects light such that the angle of incidence is equal to the angle of reflection.

We use a very low light intensity such that some time passes between the clicks of the photomultiplier. In other words, only one photon is in the experiment at any time. To prevent a photon from going straight across to the detector without being reflected, a wall is placed in the middle.

We want to investigate the reflection of light within the framework of statistical thermodynamics, except that we replace the principle of indifference with the third principle (24). We model this experiment as a two-state system. The photomultiplier P for detecting photons is one state. The universe U without P forms the second state. The classical ray model says that the mirror reflects light such that the angle of incidence is equal to the angle of reflection. Our experiment is constructed symmetrically such that a photon can only move on one path, as displayed in Figure 5. This path is the only microstate. The macrostate P has exactly one microstate, which represents this path. The macrostate U is empty. The probability of P is one and is zero for U. Unfortunately, this model disagrees with experimental results. If we cut off several parts of the mirror, including the essential middle part of the mirror, then sometimes the photomultiplier clicks. We observe reflection. Hence, the photon should also move on other paths than on the unique microstate of P.

This observation suggests claiming that all paths from the source S to P or U are the *microstates* (possibilities). Only paths through the wall are forbidden[62]. The *macrostates* (outcomes) P and U are the sets of paths from source S to P and U, respectively. The macrostates are observable. Either the detector P clicks or the photon disappears in U. Both macrostates are disjoint and cannot happen simultaneously in the present.

In thermodynamics, all accessible microstates contained in the current macrostate have, according to the principle of indifference, equal probability. This can be represented geometrically by allocating the same unit vector to every accessible microstate of a given macrostate. Summing up all these vectors yields a vector that represents the multiplicity of the macrostate. Obviously, this approach does not work. In agreement with our probabilistic framework, we replace the principle of indifference with the third principle (24): Each possible path is furnished with a reasonable amplitude as displayed in Figures 6 and 7. The phase of the amplitude is the action. In spacetime, the action is invariant with respect to Lorentz transformations[63], and thus fit into our

[62]We assume that there is only a finite number of paths to avoid complicated mathematics.
[63]Zee [1993], Chapter 7, Zwiebach [2004], Chapter 5

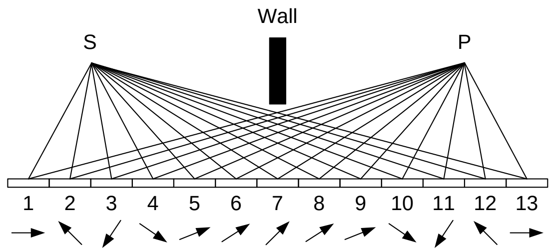


Figure 6: Feynman's view says that light has an amplitude equal in magnitude for each possible path from the source to the photomultiplier. In particular, it can be reflected from every part of the mirror, that is, from the middle as well as from the other parts.

timeless probability theory.

This is the way how the well-known *wave-particle duality* is resolved in Feynman's formulation: The photon has no complementary partner, such as a wave. Instead, paths are equipped with arrows, namely the probability amplitudes that satisfy the general rule: For each path from the source to the photomultiplier, draw an appropriate arrow, and add all arrows with parallelogram addition. This is the quantum rule of *superposition*. Then square the magnitude of the resulting arrow. This returns the probability of being detected by the photomultiplier. The wave turns out to be simply a probability distribution.

The action for photons depends only on the length of the paths. This is a geometrical property of the experimental set-up and thus belongs to the category structure.

In many textbooks, it is stated that a photon should simultaneously move on all possible paths from the source to the detector, a strange visualization[64]. This seems to be, however, the only consistent conclusion if we assume that we live in four-dimensional space-time. In our framework, this weird imagination does not apply due to our categorization trinity of time and structure. There, the amplitudes are related to possibilities, not to interactions of particles in the present.

The small arrows in this sum are displayed[65] in Figures 6 and 7 where we have divided the mirror into little squares, with one path for each square. When we add all contributions for the paths, then, as seen in Figure 7, the final arrow length evolves mainly from arrows of the middle part of the mirror, whereas the contributions from the left and right parts almost cancel out each other. This sum is the complex multiplicity (63) for the macrostate P.

More precisely, for all paths from the source to the photomultiplier, the action S_{path} is very large compared to Planck's constant. Therefore, for nearby paths, the amplitudes differ very much, since a relatively small change of the action is large compared to $\hbar$, thus yielding a completely different phase. This

[64]The usual interpretation of this experiment can be found in the nice talk of Girvin in the KITP Public Lectures, see online `kitp.edu/online/plecture/girvin`.

[65]The figures in this section are modifications of related ones in the book Feynman [1985]

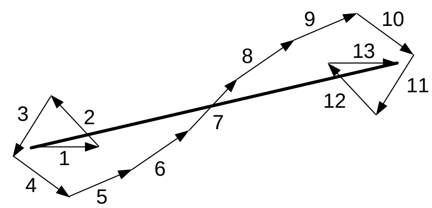


Figure 7: The amplitudes for all possible paths are added together. The major contribution to the final arrow's length is made by the paths of minimal action, that corresponds to the paths of minimal length.

implies the cancellation of the arrows in the sum.

There is only one exception, namely the paths that are infinitesimally close to the path of least action, also called the extremal path. In this case, the first variation of the action is zero. This implies that nearby paths have almost equal action and thus have equal amplitudes in the first approximation. Exactly these paths are the important ones and contribute coherently. This occurs in the region where the arrows almost point in the same direction.

In other words, all paths distant from the classical path of least action interfere destructively. On the other hand, the paths in the neighborhood of the classical path interfere constructively. This is the reason why we observe mainly classical events, such as light travels in a straight line. Only the middle part of the mirror seems to be responsible for reflections. It is astonishing, however, that the "stopwatch" $e^{\frac{i}{\hbar}S(\mu)}$ rotates ten thousands of times until the photon reaches the photomultiplier, but the amplitude for this event is the final hand direction of the watch.

This, however, is not the whole story. The fundamental question is: How does the photon find the path of extremal action? Does the photon smell out all possible paths to find the right path. Or is this approach only a mathematical description far away from any reality? If this formalism has any validity, we should be able to show in an experiment that a photon sometimes chooses also other paths.

It is simple to answer these questions using the following experiment. We cut off a large part of the mirror such that only three segments on the left side are leftover, see Figure 8. Moreover, the amplitudes are displayed in greater detail. If we add all arrows, we see that they cancel out, and the probability of being detected in the photomultiplier is almost zero.

But if Feynman's theory is true, then photons should be detected when we reduce the left part of the mirror in a manner such that no cancellation can occur, see Figure 9. Then the majority of arrows points to the right, and in total, we obtain an amplitude that predicts a strong reflection. In fact, in agreement with our theory, the photomultiplier clicks sometimes. **This sounds crazy: in theory as well as in practice, you cut off the critical middle of the mirror, from the remaining part, you scrape away appropriate pieces, and then you observe reflection. Once more, the photon seems to walk on each possible path with a stopwatch. This**

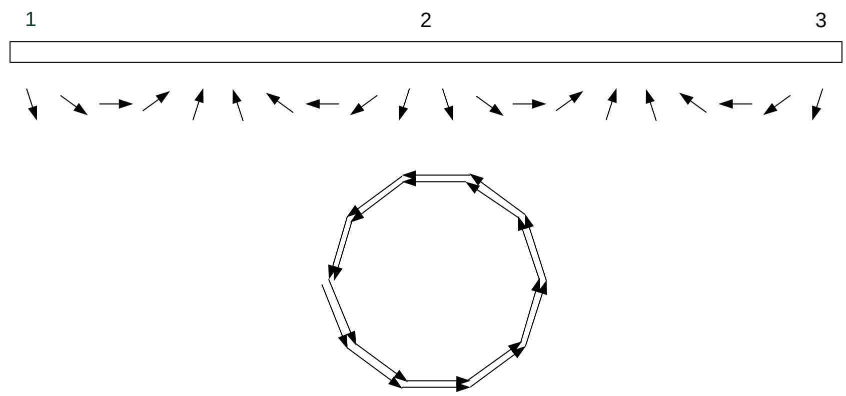


Figure 8: Considering only the piece of the left part of the mirror, the detector does not click, since the amplitudes add up to approximately zero.

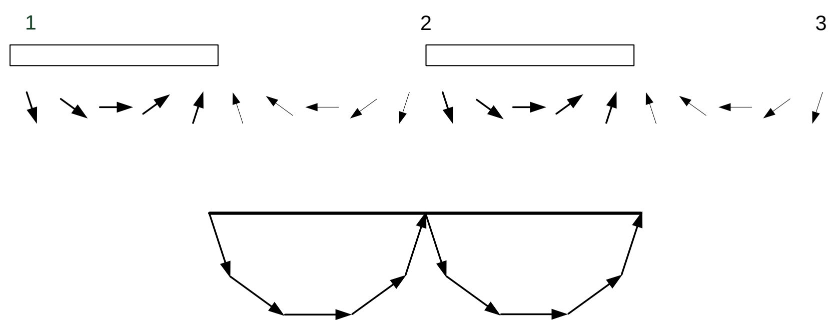


Figure 9: A striped mirror reflects a substantial amount of light, and is called a diffraction grating.

weird view vanishes when we ascribe the amplitudes as a structural concept.

The size of the experiment, the placements of source, photomultiplier and cut-out's of the mirror, and hence the direction of the arrows also depend on the color of the light, hence on its energy or equivalently, its frequency. This follows from the definition of the action that depends on the frequency or equivalently on the wavelength of light.

We have seen that cutting off parts of the mirror results in the complex multiplicity being close to zero since the amplitudes for the microstates cancel each other out. Further appropriate cutting off of the mirror then leads to amplitudes that are different from zero because no cancellation takes place. In classical thermodynamics, such situations cannot happen because all microstates are equipped with one. The second law of thermodynamics, stated in Section 5.1, does not help since it says that removing any constraint of an isolated combined system will increase the multiplicity and entropy. The reflection of light shows that for complex multiplicities and complex entropy, a more general, almost trivial law does hold: Changing constraints of an isolated combined system changes the quantum Boltzmann entropy and the complex multiplicity, thus Feynman's path integral.

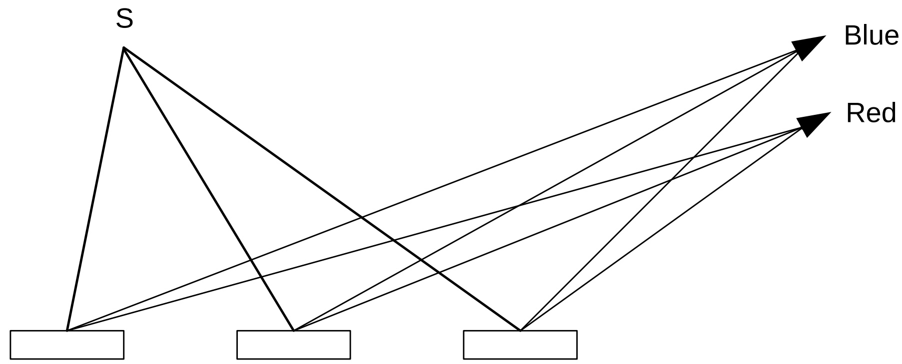


Figure 10: The angle of reflection depends on the color (wavelength) of light.

6 Quantum Information Theory

In classical information theory, information-processing tasks like storage and transmission of information, data compression, Shannon entropy, or channel capacity are investigated. This theory is based on the observation that there is a fundamental link between probability and information. Quantum mechanics is a probabilistic theory, thus connected with information and leading to quantum information theory. This is a rather new and rapidly developing discipline. The special probabilistic rules of quantum mechanics lead to fundamental differences between quantum and classical information theory. We can present only a concise overview, including some historical remarks. There are many good introductions to quantum information theory. We mention only few textbooks[66].

6.1 History of "It from BIT"

Besides the striking concepts *superposition* and *entanglement*, there are many further fundamental questions in physics. Among them: Which of the variables or quantities are basic, and which of the quantities emerge? Is time an illusion, or emerges time from a timeless physical model? Emerges even Minkowski spacetime? Is the asymmetry of time an accident? Why is the distinction between the past, present, and future almost absent in the fundamental physical models? Is at least our three-dimensional Euclidean space fundamental? Is causality violated in physics? Can we reconstruct physics by using only simple principles based on our experience? Discussions and answers to these and many similar questions cannot be found in usual textbooks. However, they are discussed in the literature. We have already considered some of these questions in terms of our categorization.

In 1990, Wheeler[67] discussed the fundamental relationship between physics, quantum theory, and information? Section 19.2 has the title " 'It from Bit' as Guide in Search for Link Connecting Physics, Quantum, and Information".

[66] Nielsen, Chuang [2010], Lyre [2004], von Weizsäcker [2006], Jansson [2017]
[67] Wheeler [1990]

He writes:

> *'It from bit' symbolizes the idea that every item of the physical world has at the bottom - a very deep bottom, in most instances - an immaterial source and explanation; that which we call reality arises in the last analysis from the posing of yes-or-no questions and the registering of equipment-evoked responses; in short, that all things physical are information-theoretic in origin and that this is a participatory universe.* John A. Wheeler, 1990

So, what is information? In 1948, Shannon[68] formulated his groundbreaking paper "The Mathematical Theory of Communication". There, he introduced a measure of uncertainty, describing an average amount of information, which he called *entropy*. John von Neumann suggested this name to Shannon with the following words:

> *"You should call it entropy, for two reasons. In the first place, your uncertainty function has been used in statistical mechanics under that name. In the second place, and more importantly, no one knows what entropy really is, so in a debate you will always have the advantage."* John von Neumann[69]

In 2015 under the name "Simons Collaboration on Quantum Fields, Gravity, and Information[70]" some of the leading researchers try to foster communication, education, and collaboration in the mentioned areas. On page 1, they write:

> *When Shannon formulated his groundbreaking theory of information in 1948, he did not know what to call its central quantity, a measure of uncertainty. It was von Neumann who recognized Shannon's formula from statistical physics and suggested the name entropy. This was but the first in a series of remarkable connections between physics and information theory. Later, tantalizing hints from the study of quantum fields and gravity, such as the Bekenstein-Hawking formula for the entropy of a black hole, inspired Wheeler's famous 1990 exhortation to derive "it from bit", a three-syllable manifesto asserting that, to properly unify the geometry of general relativity with the indeterminacy of quantum mechanics, it would be necessary to inject fundamentally new ideas from information theory. Wheeler's vision was sound, but it came twenty-five years early. Only now is it coming to fruition, with the twist that classical bits have given way to the qubits of quantum information theory.*

[68] Shannon [1948]
[69] See Tribus, McIrving [1971]
[70] See Simon [2015]

The members of Simon's collaboration include well-known leaders in quantum information and the fundamentals of physics, among them Aaronson, Aharonov, Hayden, Preskill, and Susskind.

Forgotten in Wheeler's paper [71] and Simon's proposal is C.F. von Weizsäcker[72] who dealt already in the fifties with such fundamental questions in probability, information, and physics. Concerning Simon's statement that "Wheeler's vision was sound, but it came twenty-five years early", Weizsäcker's theory came sixty years early. He assumed that quantum theory is the fundamental theory describing nature. He attempted to reconstruct this theory with binary alternatives by realizing the Kantian idea of justifying the fundamental laws of nature from our experience.

In his program, all physical objects and their properties shall be deduced from abstract quantum information theory, based on binary alternatives, nowadays called qubits. He used the name "ur" instead of "qubit" and called his theory "ur theory". The name qubit was introduced much later in 1995 and is attributed to Benjamin Schumacher. Frequently, physicists speak of spinors instead of urs or qubits. In Weizsäcker's program, a temporal logic using the structure of past, present, and future is incorporated. Already at an early stage, reconstructions of quantum theory, relativity theory, and quantum field theory were published[73].

By the way, Wheeler was invited already in 1980 by Weizsäcker to the fourth conference on Ur Theory and consequences. Wheeler spoke about "The Elementary Quantum Act as Higgledy-Piggledy Building Mechanism," and ten years later[74], Wheeler gave the lecture with the very intuitive title "It from Bit". However, Wheeler did not reference the existing research of C. F. von Weizsäcker and his co-workers[75].

6.2 Physics and Information

After these historical remarks, we ask: What is information? It is a concept associated with different phenomena, such as meaning, truth, communication, knowledge, reference, entropy, data compression, and physical processes. This word derives from the Latin verb "informare" with the meaning like "to instruct" or "to give form". An accurate definition of any fundamental concept seems to be hard, if not hopeless. A precise answer must use more fundamental concepts, which at some point cannot be further explained. Perhaps a working hypothesis might be:

- *Information* **is a quantitative measure of form**. Form is neither matter nor mind, but a property of material objects such as experimental set-ups or machines. These have form, mathematics allows to describe this form, and the form is expressed through experimental results.

[71] Wheeler [1990]

[72] See von Weizsäcker [1955], von Weizsäcker [1988], von Weizsäcker [1992], von Weizsäcker [2006], and the literature referenced therein.

[73] von Weizsäcker [1988], Görnitz, Graudenz, von Weizsäcker [1992]

[74] Wheeler [1990]

[75] See Görnitz [2019]

There is, however, an actual debate whether information is physical, whether information is more fundamental than matter, whether all things are reducible to information, or whether the universe is a computational system like a Turing machine. In the following, we briefly present the opinions of three prominent researchers.

Very early in the fifties, Weizsäcker[76] formulated some principles of his *ur theory*, see Section 5.1. He postulated in his *principle of alternatives* that physics is best formulated based on empirical decidable alternatives. Moreover, he stated in his *ur hypothesis* that all alternatives can be constructed from binary alternatives. Hence, physics should be defined entirely in terms of binary alternatives and their symmetries[77]. In other words, Weizsäcker developed the view that "all physical models can be derived from the information contained in an ur".

A *bit* is a physical quantity that can occupy one of two distinct classical states, conventionally labeled by the binary values 0 or 1. An *ur* or *qubit* is represented by a vector in a two-dimensional complex Hilbert space. It can be characterized as one *bit* of potential information. The two binary values represent two orthonormal vectors in this Hibert space. In ur theory, urs permit a decomposition of state spaces into the tensor product of two-dimensional complex Hilbert spaces. Consequently, every physical object can be described as a composition of urs, or qubits if you like. His ur theory allows an entirely new perspective on the three entities matter, energy, and gravitation. Werner Heisenberg wrote about his concept "that the realization of Weizsäcker's program requires thinking of about such a high degree of abstraction that up to now - at least in physics - has never happened." Not surprisingly, Weizsäcker's approach was hardly appreciated, perhaps it was far too abstract. Moreover, his predictions were beyond the imagination of most physicists. For instance, that one proton is made up of 10^{40} qubits is hard to believe, even today. However, a quantum field theory, particles, and a cosmological model has been presented in Weizsäcker's framework, already in the last century. His work is hardly mentioned in the literature, and not surprisingly, it cannot be found in "Simons Collaboration'[78]". His ur theory uses symmetry from the very beginning. It can be viewed as the start of a quantum theory of information, where symmetry groups are considered to give rise to the structure of space and time.

Wheeler[79] is an advocate viewing physics as information:

> *It from bit symbolizes the idea that every item of the physical world has at bottom - at a very deep bottom, in most instances - an immaterial source and explanation; that what we call reality arises in the last analysis from the posing of yes-no questions and the registering of equipment-evoked responses; in short, that all things physical are information-theoretic in origin and this is a participatory universe.*

[76]von Weizsäcker [1955],von Weizsäcker [1958]
[77]See also von Weizsäcker [1988], von Weizsäcker [1992], von Weizsäcker [2006]
[78]See Simon [2015]
[79]Wheeler [1990] page 311

> *Three examples may illustrate the theme of it from bit. First, the photon. With polarizer over the distant source and analyzer of polarization over the photodetector under watch, we ask the yes or no question, "Did the counter register a click during the specified second?" If yes, we often say, "A photon did it." We know perfectly well that the photon existed neither before the emission nor after the detection. However, we also have to recognize that any talk of the photon "existing" during the intermediate period is only a blown-up version of the raw fact, a count. The yes or no that is recorded constitutes an unsplittable bit of information.* Wheeler 1990

Continuing, Wheeler[80] says:

> *To the question, "How come the quantum?" we thus answer, "Because what we call existence is an information-theoretic entity." But how come existence? Its as bits, yes; and physics as information, yes;* Wheeler 1990

Not surprisingly, Wheeler's understanding is very closely related to Weizsäcker's framework. This is expressed in the same paper in Section 19.6:

> *19.6 Agenda*
> *Intimidating though the problem of existence continues to be, the theme of it from bit breaks it down into six issues that invite exploration:*
> *One: Go beyond Wootters and determine what, if anything, has to be added to distinguishability and complementarity to obtain all of standard quantum theory.*
> *Two: Translate the quantum versions of string theory and of Einstein's geometrodynamics from the language of continuum to the language of bits. ....* Wheeler 1990

We remark that Wootters[81] derived 1980 the Hilbert space together with its complex probability amplitudes mainly from the fundamental demands of complementarity and distinguishability.

The contrary view is that information depends on physical objects or systems, rather than the other way around. In other words, experimental set-ups or machines have form and contain all information. A prominent advocate of this view is Preskill[82], who writes:

> *Information, after all, is something that is encoded in the state of a physical system; computation is something that can be carried out on an actual physically realizable device. So the study of information and computation should be linked to the study of the underlying physical processes. [....] "The moral we draw is that 'information is physical'.* Preskill 1998

[80] Wheeler [1990] page 313
[81] Wootters [1981],Wootters [1981]
[82] Preskill [1998], pp. 7, 10

It means that information only exists when encoded in any physical device. This is supported by the fact that the device has form. For example, an electron in a double-slit experiment depends not only on the physics of the electron but on each detail of the experiment.

We have now come to completely contradictory views of information and physics. But neither of the theses, whether information can be reduced to physics or vice versa, can be falsified by mathematics or experiments. Perhaps, each of these conceptions has its justification.

Today, Weizsäcker's ur theory and his ur hypothesis, stating that all alternatives can be constructed from binary alternatives, form the basis of quantum information. He can be called the founding father of quantum information theory.

For an intuitive understanding of the concept of physical information, we imagine a physical object, such as a coin, a door, or an electron, and we ask: What is the information content of this physical object? Think of some people who share the same background about this object, but they don't know the actual state: heads or tails, a door is open or closed, an electron has spin up or down. It is natural to define the *amount of information* of the object as the alternatives or instructions which are necessary to be able to reconstruct the state of the object. Obviously, for our three objects, we need only one binary alternative, namely a *bit*, to identify in each case the state. A bit can be viewed as a question with two possible answers. The amount of information of a Bescon die with 4 facets can be reduced to two bits, namely whether we get the lower numbers 1 and 2 or not and whether we get an odd number or not. In this sense **information is physical**, or we can say that **physical objects carry information.** It is hard to believe that information and logic would exist in an empty universe without any objects. Information can be not a purely mathematical idea but is dictated by the substance in our universe. We can possibly say that **information without physical objects is just as meaningless as physical objects that cannot be described by information.**

Given an n-fold alternative $X = \{x_1, ..., x_n\}$ describing an object with n mutually exclusive states. What is its amount of information? If X consists of only one element, then no question is necessary to obtain a state, and the amount of information is zero. If X consists of two elements, it is a *bit*, only one question is required for getting a state, and the amount of information is one. If X consists of four elements, then two questions are required, and the amount of information is two. If X consists of three elements, then either one or two questions are required, and the amount of information is defined as $I = \log_2 3 = 1.585$. If X consists of n elements, then the *amount of information* I is defined in terms of binary questions as

$$I = \log_2 n. \tag{67}$$

This number corresponds to the height of the related binary decision tree, that is, the largest number of edges in a path from the root node to a leaf node. Each edge in the tree represents a binary decision with values 0 or 1. Hence, a path corresponds to a binary *register*. The amount of information is a characteristic number of this tree and belongs to the category structure of physical objects.

6.3 Information and Shannon-Entropy

Shannon[83] was interested in his groundbreaking paper "The mathematical Theory of Communication" in a theory describing the communication of information, not in information itself. In his theory an alternative X, called the *source*, generates consecutively random outcomes x_k, called *letters*. The outcomes obey the probability distribution $p_k = \Pr(x_k)$. Shannon asked:

- What is a *measure of information* or the *uncertainty* for the next outcome x_k associated with a probability distribution p_k?

In contrast to the geometrical height (67) of a binary register, this is a probabilistic question.

Shannon was not interested in semantic aspects and the meaning of information. His starting problem was to reproduce at least approximately a message that arrived from another point. The simplest case of a probability distribution is that all outcomes are equally likely. Then the probability for each of n outcomes is $p_k = 1/n$. Think of throwing a coin 4 times. The sample space consists of all quadruples with values 0 or 1 (Heads or Tails), thus contains $n = 2^4$ elements. The amount of information (67) of one element is $\log_2 2^4 = 4$, and the probability is $1/16 = 1/(2^4)$. We have to multiply the probabilities 1/2 for four questions to get the probability of one outcome. Hence, it is natural to assign to the uncertainty the number 4. Thus, the amount of information coincides with the uncertainty for equally likely events.

More general, the identity

$$I = \log_2 n = -\log_2 1/n = -\log_2 p_k \tag{68}$$

indicates to define for equally likely probabilities the uncertainty as the negative logarithm of the probability. **In this case, both values, the deterministic amount of information (67) and the uncertainty associated with a probability distribution, are equal although they have completely different derivations and interpretations**. One represents the characteristic number of the related decision tree; the other one represents a random walk through this tree.

What is the average uncertainty of X for equally likely outcomes? This is just the average value

$$\sum_{k=1}^{n} \frac{1}{n} \log_2 n = -\sum_{k=1}^{n} p_k \log_2 p_k = \log_2 n, \tag{69}$$

where $p_k = 1/n$. Of course, this value coincides with the amount of information, that is, the number of binary decisions necessary to specify any outcome. Hence, we have a mathematical identity, not a physical one.

[83]Shannon [1948]

Now let us look at general non-uniform probabilities p_k? This case corresponds to binary questions, which may not be equally likely. What is the average uncertainty of X? The most reasonable generalization of formula (69) is just the average value

$$S_S = S_S(X) = -\sum_k p_k \log_2 p_k. \tag{70}$$

The quantity S_S is called the *Shannon entropy* or *information entropy*. We speak of this entropy as a *measure of missing information* associated with any probability distribution, or simply of a measure of *uncertainty*? It is formally identical with the Gibbs entropy and the von Neumann entropy (61).

The entropy S_S is (i) a monotonically decreasing function of the probabilities p_k, (ii) is additive for independent events, where probabilities are multiplied, (iii) does not change if events with zero probability are added, and (iv) is maximized when all probabilities are equal. It is well-known that S_S is, up to a constant k, the unique function satisfying these four criteria.

6.4 Data Compression

As an application of entropy, let us consider *data compression*. Suppose we have a data file, say a text containing a sequence of letters. Depending on the text, different letters occur with different frequencies. In data compression, we want to compress the given text into the smallest possible data file. What is a good compression algorithm?

It is standard to write an uncompressed text as numbers in ASCII code where every letter is assigned a number represented with 8 bits. In books and other texts, some letters, such as 'e', 'a' or 'o', are much more common than 'x', 'q' or 'z'. Hence, it should be possible to compress texts rather efficiently. Let us look at the short text X consisting of three words "begin prologue end". The text consists of 16 letters with corresponding frequencies in brackets: e(3), g(2), n(2), o(2), b(1), i(1), p(1), r(1), l(1), n(1), d(1). We expect that a good compression algorithm would use fewer than 8 bits for the letters with high frequency. Our text has 11 different letters. Hence, the most straightforward coding requires 4 bits, since $11 < 2^4$.

But what is the smallest number of bits to encode a given text? Shannon's answer was: this is the entropy S_S. He proved that, on average, it is not possible to encode data with fewer than S_S bits. This is the well-known *source coding theorem*. For our text we obtain

$$S_S = -\left(\frac{3}{16}\log_2\frac{3}{16} + 3\cdot\frac{1}{8}\log_2\frac{1}{8} + 7\cdot\frac{1}{16}\log_2\frac{1}{16}\right) = 3.328. \tag{71}$$

Shannon's theorem states that on average the best we can do is to represent the letters above with 3.328 bits. A text with 1000 letters and the same entropy could be encoded with 3328 bits. If we would use 4 bits for each letter, we need 4000 bits. A good algorithm uses the extra information of different probabilities.

When we use words rather than letters, the entropy decreases. For our text with three words, assumed to be uniformly distributed, we obtain $S_S = -3 \cdot 1/3 \log_2 1/3 = 1.585$. This leads to a compression factor of $4/1.585 = 2.523$. We refer the reader for compression algorithms to the literature but mention that many file systems automatically compress files when storing them. Images on the web are compressed, frequently in a JPEG or GIF format.

6.5 Von Weizsäcker's Reconstruction of Physics

Already 1955 von Weizsäcker introduced a quantum theory of information which he called *ur theory* [84]. His theory is a consistent interpretation of quantum theory in terms of information. It is the quantum theory of empirical decidable alternatives. It is based on his *ur hypothesis* which says that

- (i) all physical objects, quantities, and all dynamics are characterized by alternatives, and (ii) all alternatives can be constructed from binary alternatives.

In other words, he postulated that quantum theory should be taken as the basic theory for all physical models. This principle has fundamental consequences. Some of them are discussed in the following.

U(2) Symmetry and the Einstein Space

In Weizsäcker's ur theory, dynamics is modeled as interactions between urs, that is, qubits which are vectors in $\mathbb{C}^2$ representing binary alternatives. The continuous symmetry group of a binary alternative is the unitary group $U(2)$. It contains two subgroups, namely the special unitary symmetry group $SU(2)$ and the circle group $U(1)$. It follows that nothing changes if all urs are simultaneously transformed with the same unitary matrix in $U(2)$.

A fundamental question is to understand the three-dimensional position space and the four-dimensional spacetime. If all physical objects, including particles and spaces, are constructed via binary alternatives, then answers should come off from the symmetry group of urs. Mathematically, $SU(2)$ is locally isomorphic to the three-dimensional rotation group $SO(3)$ in Euclidean space. Thus, it is natural to view it as the rotation group in a three-dimensional real position space or in a three-dimensional real momentum space. In ur theory, the group $U(1)$ is interpreted as the group of the temporal changes. Therefore, spacetime emerges from the invariance group of the ur.

At the moment we have derived a three-dimensional reference frame and a further coordinate, both may be interpreted as position and time. It is interesting to look at the unitary matrices in $SU(2)$ itself. It is easy to show that each unitary matrix U takes the form

$$U = \begin{pmatrix} w + iz & ix + y \\ ix - y & w - iz \end{pmatrix}, \tag{72}$$

[84] See Section 1.7 von Weizsäcker [1992], von Weizsäcker [2006], von Weizsäcker [1955]

where the real coordinates x, y, z, w satisfy the normalization condition

$$x^2 + y^2 + z^2 + w^2 = 1. \tag{73}$$

These four coordinates, together with the normalization condition, are the points of the so-called *Einstein space*. It is the three-dimensional sphere $\mathbb{S}^3$ in a four-dimensional real space. In the theory of general relativity, the Einstein space is a special solution of Einstein's field equations[85]. Therefore, the symmetry group of an ur corresponds uniquely to the sphere $\mathbb{S}^3$. Moreover, we have a second derivation of how spacetime emerges from the invariance group of urs. This derivation says that each state of an ur can be mapped to a point on the $\mathbb{S}^3$, leading to the surprising observation that the topology of the cosmos would result directly from quantum theory. Many physicists have a different point of view. They think of gravity as an emergent phenomenon that arises from collective statistical behavior.

Weizsäcker thought about the essential question: What must be added to distinguishability and complementarity to obtain the rules of quantum theory? He published his reconstruction program in the article "Komplementarität und Logik[86]", which is dedicated to Bohr's 70th birthday. A recommendable survey on his program for reconstruction is written by Görnitz and Ischebeck[87]. In summary, Weizsäcker attempted to unify physics, rather than to give another interpretation.

Binary Alternatives

Physics investigates decidable alternatives. An alternative represents attributes or properties of any physical object or experiment. In classical logic an *n-fold alternative*

$$a = \{a_1, ..., a_n\}\,,\ n \in \mathbb{N} \tag{74}$$

is a set of n mutually exclusive statements or possibilities a_i, where exactly one turns out to be true if an empirical test happens, but none of the other a_j with $j \neq i$. A *binary alternative* is a 2-fold alternative.

It turns out that information and knowledge in physics can be ascribed to n-fold alternatives via binary alternatives. From the point of view of logic, this is evident, since every n-fold alternative can be decomposed by deciding k binary YES-NO questions successively, where $2^k \geq n$. Thus, we get a logical decomposition of an alternative into a set of binary alternatives.

However, a decomposition of an alternative into binary alternatives is not unique. For example, the 4-fold alternative consisting of the possibilities a_1, a_2, a_3, a_4 can be decomposed in several different ways. Let us look at the two binary alternatives b and c:

$$b_1 = \{a_1, a_2\}\,,\ b_2 = \{a_3, a_4\}\,, \tag{75}$$

[85] See von Weizsäcker [2006] page 121
[86] von Weizsäcker [1955],von Weizsäcker [1958]
[87] Görnitz, Ischebeck [2003]

and

$$c_1 = \{a_1, a_3\}\,,\; c_2 = \{a_2, a_4\}\,. \tag{76}$$

Then the possibilities of a can be written as

$$a_1 = \{b_1, c_1\}\,,\; a_2 = \{b_1, c_2\}\,,\; a_3 = \{b_2, c_1\}\,,\; a_4 = \{b_2, c_2\}\,. \tag{77}$$

We can define a third binary alternative d

$$d_1 = \{a_1, a_4\}\,,\; d_2 = \{a_2, a_3\}\,, \tag{78}$$

and describe a in terms of the binary alternatives b and d, yielding another decomposition

$$a_1 = \{b_1, d_1\}\,,\; a_2 = \{b_1, d_2\}\,,\; a_3 = \{b_2, d_2\}\,,\; a_4 = \{b_2, d_1\}\,. \tag{79}$$

If we assign the logical values 0 and 1 to the binary alternatives, we obtain a *register representation* of alternatives. For example, when assigning 0 to b_1 and c_1, and 1 to b_2 and c_2, then we get the registers

$$a_1 = 00,\; a_2 = 01,\; a_3 = 10,\; a_4 = 11. \tag{80}$$

For simplicity we have dropped the brackets.

Thus, from the logical point of view, the decomposition of an alternative into binary alternatives is almost trivial. How can we decompose alternatives in the *vector representation*? We assign the values 0 and 1 to orthonormal unit vectors in $\mathbb{C}^2$:

$$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix},\; |1\rangle = \begin{pmatrix}0\\1\end{pmatrix}. \tag{81}$$

We use the well-known *Dirac's "bracket" notation.* Weizsäcker called each two-dimensional complex vector *ur*. We can rewrite (80) as tensor products of vectors:

$$\begin{aligned}
a_1 = |00\rangle &= \tbinom{1}{0} \otimes \tbinom{1}{0} = (1\;0\;0\;0)^T,\\
a_2 = |01\rangle &= \tbinom{1}{0} \otimes \tbinom{0}{1} = (0\;1\;0\;0)^T,\\
a_3 = |10\rangle &= \tbinom{0}{1} \otimes \tbinom{1}{0} = (0\;0\;1\;0)^T,\\
a_4 = |11\rangle &= \tbinom{0}{1} \otimes \tbinom{0}{1} = (0\;0\;0\;1)^T.
\end{aligned} \tag{82}$$

The alternatives a_i are represented as orthonormal vectors contained in the tensor product $\mathbb{C}^4 = \mathbb{C}^2 \otimes \mathbb{C}^2$. Hence, we speak also of *base states*[88].

Quantum theory describes observable values in terms of self-adjoint operators with a discrete eigenvalue spectrum and an orthonormal basis. Because of the orthonormality condition such an operator can be interpreted as an empirically decidable alternative.

[88]See Section 4.2 Jansson [2017] for further discussions of vector and register representations.

It is easy to see that each n-dimensional Hilbert space can be embedded into a subspace of the tensor product of k two-dimensional Hilbert spaces

$$\mathbb{C}^n \subseteq \bigotimes_k \mathbb{C}^2,\ 2^k \geq n, \tag{83}$$

such that the linear structure and the metric are maintained.

The resulting quantum theory is obtained by assigning to the alternative orthonormal basis vectors, say $|a_i\rangle$ in an n-dimensional complex Hilbert space $\mathbb{C}^n$. Then an n-fold alternative is identified with some vector $|a\rangle \in \mathbb{C}^n$, the latter embedded in the tensor product $\bigotimes_k \mathbb{C}^2$. Thus, the ur hypothesis is logically and mathematically correct.

The binary alternatives b, c and d are distinct, leading to urs that are *distinguishable*. We can speak of the *Boltzmann-statistic* of urs.

Bosonic Representations

Each n-fold alternative $a = \{a_1, ..., a_n\}$ can be decomposed also into $n-1$ *indistinguishable urs*. Let the binary alternative e be defined as

$$e_0: \text{ index i of } a \text{ stays}, \quad e_1: \text{ index i of } a \text{ is replaced by i+1}. \tag{84}$$

For $n = 4$ we get the decomposition

$$a_1 = \{e_0, e_0, e_0\},\ a_2 = \{e_0, e_0, e_1\},\ a_3 = \{e_0, e_1 e_1\},\ a_4 = \{e_1, e_1, e_1\}. \tag{85}$$

Vectorization leads to the base states in $\mathbb{C}^8$:

$$\begin{aligned} a_1 = |000\rangle &= \binom{1}{0} \otimes \binom{1}{0} \otimes \binom{1}{0}, \\ a_2 = |001\rangle &= \binom{1}{0} \otimes \binom{1}{0} \otimes \binom{0}{1}, \\ a_3 = |011\rangle &= \binom{1}{0} \otimes \binom{0}{1} \otimes \binom{0}{1}, \\ a_4 = |111\rangle &= \binom{0}{1} \otimes \binom{0}{1} \otimes \binom{0}{1}. \end{aligned} \tag{86}$$

Since the alternatives are invariant with respect to permutations of the binary alternative e, we get

$$\begin{aligned} &a_1 = |000\rangle, \\ &a_2 \in \{|001\rangle, |010\rangle, |110\rangle\}, \\ &a_3 \in \{|011\rangle, |110\rangle, |101\rangle\}, \\ &a_4 = |111\rangle. \end{aligned} \tag{87}$$

It is likely to represent the alternatives in the symmetric form

$$\begin{aligned} &a_2 = \frac{1}{3}(|001\rangle + |010\rangle + |110\rangle), \\ &a_3 = \frac{1}{3}(|011\rangle + |110\rangle + |101\rangle). \end{aligned} \tag{88}$$

All considerations above can be easily generalized to arbitrary n-fold alternatives. For example, consider a roulette with a ball in one of the 37 cells, describing a 37-fold alternative. Then this ball can be identified with the cell containing the ball, and the cell can be identified with a bosonic representation consisting of zeros and ones.

On Some Large Numbers in Physics

In physics, usually one works with numbers, such as natural, nonnegative, real, or complex numbers and corresponding vectors and matrices. In contrast, we have seen that the ur theory basically works with logical quantities, namely questions and decisions. For example, a fundamental problem is: how many questions are necessary to describe any single physical object?

The ur theory provides a "logical atomism" in the sense that the smallest objects are not spatially small but logically small. However, it should be noticed that an ur is by no means a particle. It's just one bit of information representing one YES-NO decision. This is a language that uses bits and sequences of bits, called *registers*. At a first glance, this leads to an unusual way of thinking. In the following, we discuss several aspects of the relationship between some cosmological numbers and the ur theory[89]. In particular, we illustrate the ur language when calculating the number of nucleons in the universe and the photon-nucleon ratio. We work with approximate results where prefactors of order 10 or 100 are neglected in the following rough estimates.

The *Compton wavelength* $\lambda = h/mc$ is known to be a measure of the size of particles. The ratio of a plausible estimated cosmic radius $R \approx 10^{26}$ m, and the Compton wavelength of a proton $\lambda_p \approx 1.3 \cdot 10^{-15}$ m is

$$E_1 = \frac{R}{\lambda_p} \approx 10^{40}. \tag{89}$$

E_1 is called the first *Eddington number*. It is the ratio of a cosmological distance and an atomic distance. Surprisingly, the ratio of the force between an electron and a proton e^2/r^2 and their gravitational force Gm_em_p/r^2 is approximately equal to E_1, such that

$$E_1 \approx \frac{e^2}{Gm_em_p} \approx 10^{40}. \tag{90}$$

Notice that this number is independent of the distance between the electron and the proton.

The second Eddington number E_2, defined as the number of nucleons in the universe and estimated according to the cosmological mass density, is about

$$E_2 \approx 10^{80}. \tag{91}$$

. Hence, it is the square of the first Eddington number.

[89]See Lyre [1995], Lyre [2004], von Weizsäcker [2006], and the references therein for further discussions.

Eddington observed that the volume with classical electron radius coincides with the product of the cosmic radius R and the Planck area l_p^2, where $l_p = (\hbar\kappa/c)^{1/2}$ is the Planck length and $\kappa = 8\pi G/c^2$ the gravitational constant. Since the Compton wavelength of a proton is almost equal to the classical electron radius, it follows that

$$\lambda_p^3 \approx R\, l_p^2, \tag{92}$$

If we divide the volume of the observable universe into three-dimensional disjoint cells of size λ_p^3 and it's surface into two-dimensional disjoint cells of the Planck area, then the Planck pixels on the surface correspond to the cells of volume λ_p^3. In other words, the three-dimensional world seems to be an image of data that can be stored on a two-dimensional projection. Now, this is known as the *holographic principle*.

The third Eddington number is defined as

$$E_3 \approx \frac{R^3}{\lambda_p^3} \approx \frac{R^2}{l_p^2} \approx 10^{120}. \tag{93}$$

Summarizing, we have obtained the relations $E_1^2 = E_2$ and $E_1^3 = E_3$. Already Dirac emphasized that all this cannot be an accident and needs an explanation. In the following, we show how ur theory explains such numbers.

A central assumption in ur theory is that all physical objects are only measurable in position space[90]. This assumption coincides with the experience in experimental physics where at the end the measurement apparatus measures a position. Theoretically, this agrees with Heisenberg's uncertainty principle, where precise knowledge of a position makes momentum completely uncertain.

It is natural to assume that the length size of a massive particle is determined by its Compton wavelength λ. The measurement of smaller distances would require high energies that destroy the particle. It follows that the cells with the smallest possible volume, which may contain a massive particle with Compton wavelength λ in the three-dimensional space, have the value λ^3. Hence,

$$N \approx R^3/\lambda^3 \tag{94}$$

is the number of disjoint cells with the smallest possible volume which may contain such a particle. This number forms an N-fold alternative that can be represented by N bosonic urs.

A binary alternative can be defined by asking whether the cell contains a particle or not. How many bosonic urs are necessary for deciding this question? For localizing a particle[91] on a line segment with radius R, we have n disjoint intervals of width λ, that is,

$$n\lambda \approx R. \tag{95}$$

[90] See Lyre [1995], page 3.

[91] We neglect the prefactor 2 as already mentioned.

An ur can describe a decision on a vertical line of positions with "UP or DOWN". Using the bosonic representation, we need $n-1$ binary alternatives, that is, $n-1$ urs for localizing this particle. We say that the *particle contains $n-1$ urs, or is identified with these urs.* In the three-dimensional space, the three spatial axes require $3(n-1)$ urs for identifying a cell of volume λ^3. Because working only with rough estimates, we identify a particle of Compton wavelength λ with n bosonic urs.

A fundamental question in physics is: How many particles are in our universe? Since nucleons (consisting of protons and neutrons with roughly the same Compton wavelength) form the ponderable matter in our universe, which is otherwise almost void, it is natural to ask for the number of nucleons. We know, this is the second Eddington number $E_2 = 10^{80}$. In the same sense, we ask: How many bosonic urs exist in our universe?

Since the Compton wavelength, λ_p of a proton is a good approximate measure for nucleons, asking for the number of bosonic urs in our universe, we consider the Compton wavelength of protons only. Then the universe can be divided simultaneously into cells with volume λ_p^3.

From (89) and (95) it follows that for detecting a proton on a line we need

$$n_p \approx 10^{40} \tag{96}$$

bosonic urs, that is, a register with about 10^{40} zeros and ones. This register can be identified with one proton; or we can also say that the proton consists of 10^{40} urs.

The cells with smallest possible volume containing a nucleon in the three-dimensional space has the value λ_p^3. Hence, the cosmos could in principle be partitioned into

$$N = R^3/\lambda_p^3 \approx 10^{120} \tag{97}$$

cells that may contain nucleons. Thus, we postulate that the total number of bosonic urs in the universe is 10^{120}. On the other hand, from (96) we know that a proton or nucleon consists of $n_p \approx 10^{40}$ urs. Therefore, if all cells accommodate nucleons, we can estimate their number by

$$m_p = N/n_p \approx 10^{80}. \tag{98}$$

This is the second Eddington number $E_2 = 10^{80}$ which was estimated empirically by measuring the cosmological mass density. The ur theory has used the Compton wavelength of a proton and the estimated cosmological radius only. Then, taking the bosonic representation of the ur theory, we can explain the second Eddington number without any further measurements. This was one of the first testable consequences of the ur theory.

Another application[92] is the determination of the *photon-nucleon ratio*, an important cosmological number. This ratio is estimated between 10^8 and 10^{10}. Similar considerations as above prove this ratio. Thus it seems reasonable to speak of confirmation of ur theory since it's hard to believe that these estimates of large numbers are an accident.

[92] See Lyre [1995], Schramm [1996] page 285.

6.6 Reconstruction of Relativity Theory

So far, we have reconstructed quantum mechanics as a timeless theory, the theory of future possibilities where nothing happens. In physics, the theory of relativity is beside quantum theory and statistical thermodynamics another fundamental theory. The center of this theory is the four-dimensional spacetime and the Lorentz transformation. Can we reconstruct this theory out of a timeless framework?

At a first glance, this seems to be impossible. However, in my lecture notes[93], a reconstruction of the mathematical formalism of special relativity based on a timeless (3+3)-position-velocity space is presented. In particular, we derived the key of relativity theory, namely the *Lorentz transform*, without any assumption about "propagation of light". It was shown that the Euclidean position-velocity space[94], 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. We refer the reader to this book.

Finally, remember that von Weizsäcker has derived spacetime, the Einstein space and relativity from $U(2)$ symmetry, see Section 6.5.

6.7 Bell's Theorem

Bell's fundamental theorem and inequalities investigate hidden-variable theories and local realism[95]. Roughly spoken, it was proved that quantum theory is incompatible with local hidden-variable theories. This research gave much insight into quantum information theory. In this final section, we examine how our probabilistic framework can be applied to related questions.

We consider the following experiment: There are two spatially separated calcite crystals, say A and B. Between them is a source that produces pairs of photons. One photon moves toward A, and the other one toward B. The pairs of photons are entangled: They are always polarized in perpendicular directions, provided both crystals have the same optical axis. The interaction of a photon with one crystal instantaneously changes the polarization of the other one. They seem to influence each other non-locally.

Einstein regarded this phenomenon as "spooky action at a distance“. With two colleagues he formulated in the well-known EPR paper[96] this paradox as follows:

> *If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of reality corresponding to that quantity.*

[93] Jansson [2017], Sections 4.13 and 4.14

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

[95] Bell [1987] Ch. 16, Sakurai [1994] pp. 174-187, 223-232, Bellac [2011] Ch.6, Jansson [2017] Section 2.13

[96] Einstein, Podolsky, Rosen [1935]

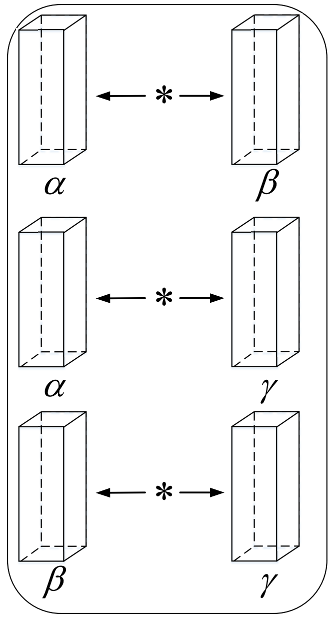


Figure 11: Three separate random experiments on pairs of entangled photons. The optical axes of the crystals are α, β and γ.

They concluded that quantum mechanics is an incomplete theory that should be extended with hidden variables. The discussion around this paper is known as the EPR paradox. EPR corresponds to the initials of the three authors.

In the case of photons, the entanglement mostly relates to polarization. However, photons can also be entangled with regard to the direction of flight.

Bell investigated the EPR paradox supposing that the photon-producing source ascribes hidden variables to each photon. We follow the approach of Sakurai (1994), who assumed:

- Assumption 1: A hidden variable is assigned that labels one photon as horizontally polarized and the other one as vertically polarized, for any optical axis.

- Assumption 2: Each photon, when interacting with a crystal, has an infinite number of hidden variables that correspond to the crystal's optical axes.

We suppose that the two crystals have three possible optical axes, which we label α, β, and γ. We perform three types of random experiments, as displayed in Figure 11.

Because of assumption (2) each photon has its own polarization state with respect to each of these optical axes. We label the states with $+$ for horizontally polarized and $-$ for vertically polarized states, respectively. For three axes, we obtain for the photon pair eight possible polarization states, see Table 2.

In the first random experiment, crystal A has optical axis α and crystal B has optical axis β. It is assumed that the polarization of the produced photon pair exist for the three optical axes α, β, and γ, although the third crystal with

Crystal A	Crystal B	Probability
α, β, γ	α, β, γ	
$+++$	$---$	Pr(1)
$++-$	$--+$	Pr(2)
$+-+$	$-+-$	Pr(3)
$+--$	$-++$	Pr(4)
$-++$	$+--$	Pr(5)
$-+-$	$+-+$	Pr(6)
$--+$	$++-$	Pr(7)
$---$	$+++$	Pr(8)

Table 2: Each row describes the polarization of a photon pair for three optical axes α, β, and γ, and their probabilities Pr(j).

optical axis γ is not part of this experiment. However, the eight possibilities are interpreted as outcomes that have the probabilities Pr(j), j=1,...,8.

The non-negative number $\Pr(\alpha = +, \beta = +)$ denotes the probability that both photons are horizontally polarized with respect to the two axes α and β, when interacting with both crystals. Since the probability of two mutually exclusive outcomes can be added, from Table 2, it follows immediately that

$$\Pr(\alpha = +, \beta = +) = \Pr(3) + \Pr(4). \tag{99}$$

Similarly, for the other two random experiments we get

$$\Pr(\alpha = +, \gamma = +) = \Pr(2) + \Pr(4), \tag{100}$$

and

$$\Pr(\gamma = +, \beta = +) = \Pr(3) + \Pr(7). \tag{101}$$

Adding together implies Bell's inequality as presented by Sakurai:

$$\Pr(\alpha = +, \beta = +) \leq \Pr(\alpha = +, \gamma = +) + \Pr(\gamma = +, \beta = +). \tag{102}$$

In the literature, it is argued that quantum theory predicts a violation in the inequality (102. This violation is experimentally verified[97].

In the following we argue that (i) the explanation of this experiment does not require quantum mechanics, (ii) classical statistical mechanics is sufficient, (iii) Bell's inequality does not describe this experiment appropriately, and (iv) our probabilistic frameworks is a nice guide for investigating this experiment.

Actually, we have not one but three different experiments, see Figure 11. Let us look at the first one with optical axes α and β. Then the pair of entangled photons has four possibilities when interacting with both crystals:

$$\alpha = +, \beta = +;\ \alpha = +, \beta = -;\ \alpha = -, \beta = +;\ \alpha = -, \beta = -. \tag{103}$$

There are no internal possibilities. The outcomes coincide with these four elementary possibilities. Hence, we have a classical probabilistic experiment.

[97]Aspect et al. [1982]

What can we say about the probability amplitudes and the probabilities? The *law of Malus*[98], already formulated in 1810, states that the intensity of a beam of light that has passed two polarizers with optical axes α and β is proportional to $\cos^2(\beta - \alpha)$. If only one photon is in the experiment, just as it is possible today in the experiments, the intensity should be identified with the probability. Thus, the probability is a function of the angles between the optical axes.

In the case $\beta = \alpha$, the probabilities

$$\Pr(\alpha = +, \beta = +) = \Pr(\alpha = -, \beta = -) = 0 \tag{104}$$

since the entangled pair of photons is always polarized at right angles. Vice versa, if $\beta = \alpha \pm \pi/2$ the probabilities

$$\Pr(\alpha = +, \beta = -) = \Pr(\alpha = -, \beta = +) = 1/2. \tag{105}$$

Generally, the law of Malus suggests

$$\Pr(\alpha = +, \beta = +) = \Pr(\alpha = -, \beta = -) = \frac{1}{2}\sin^2(\beta - \alpha), \tag{106}$$

and

$$\Pr(\alpha = +, \beta = -) = \Pr(\alpha = -, \beta = +) = \frac{1}{2}\cos^2(\beta - \alpha). \tag{107}$$

These four numbers are non-negative and adding them up gives one. Indeed, these probabilities are experimentally verified. The related probability amplitudes are their roots. This finishes the application of our probabilistic framework.

What happens with Bell's inequality (102) if we insert these probabilities? Without loss of generality, we set $\alpha = 0$. Then

$$\frac{1}{2}\sin^2(\beta) \leq \frac{1}{2}\sin^2(\gamma) + \frac{1}{2}\sin^2(\gamma - \beta). \tag{108}$$

This inequality must be fulfilled for all angles β and γ. For the angles $\beta = 3\gamma$ the inequality (108) becomes

$$0 \leq \frac{1}{2}\sin^2(\gamma) + \frac{1}{2}\sin^2(-2\gamma) - \frac{1}{2}\sin^2(3\gamma). \tag{109}$$

The function on the left hand side is negative for the angles γ between 0.1 and 0.5. Hence, Bell's Theorem is violated. What are the reasons for this violation?

The first reason is assumption 2, which assigns infinitely many optical axes to the tiny photon, and leave the large crystal completely out of consideration. Our approach is exactly the opposite: the crystal is the machine that is characterized by all possibilities, while the photon only interacts and chooses exactly one possibility. The experimental set-up determines all probabilities.

[98] Halliday, Resnick, Walker [2005]

The latter belong to the category future. The interaction is part of the present only.

The second reason is the incorrect description of the experiment, where the three experiments are pretended to be one. For example, in the experiment with optical axes α and β the value $+++$ denotes an outcome or elementary event of the experiment. But there was no interaction of a photon with some crystal with optical axis γ. Hence, this triple value is not an outcome and thus, has no probability as assumed in Bell's inequality.

This experiment also has little to do with quantum mechanics. We have described it classically with the old law of Malus from 1810, long before quantum mechanics started. The typical quantum superposition did not appear in our derivation.

References

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.

P. Attard. *Thermodynamics and Statistical Mechanics: Equilibrium by Entropy Maximisation*, Academic Press, 2002.

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

J.J. Bell. *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, 1987.

M. Le Bellac. *Quantum physics*, Cambridge University Press, 2011.

A. Ben-Naim. *Time's Arrow (?) The Timeless Nature of Entropy and the Second Law of Thermodynamics.*, Princeton University Press, 2018.

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

R.P. Crease. *The Most Beautiful Experiment*, Phys. World **15**(9): 19–20, 2002.

A. Eddington. *The Nature of the Physical World:*, THE GIFFORD LECTURES 1927. Vol. 23, BoD-Books on Demand, 2019.

A. Einstein, B. Podolski, N. Rosen, . *Can quantum-mechanical description of physical reality be considered complete?*, Physical review 47.10 (1935): 777, 1935.

R.P. Feynman, *Space-Time Approach to Non-Relativistic Quantum Mechanics* , Rev. of Mod. Phys., Vol. 20 Nr.2, 367, 1948.

R.P. Feynman, R.B. Leighton, and M. Sands. *The Feynman Lectures on Physics, Vol. III*, California Institute of Technology, 1963, Addison Wesley; Later Printing edition, 1971, `http://www.feynmanlectures.caltech.edu/`.

R. P. Feynman, A. R. Hibbs. *Path Integrals and Quantum Mechanics*, McGraw, New York, 1965.

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

C.A. Fuchs. *Quantum Mechanics as Quantum Information (and only a Little More)*, arXiv preprint quant-ph/02050392002. 2002.

C.A. Fuchs, and A. Peres. *Quantum Theory Needs No 'Interpretation'*, American Institute of Physics, 2000, S-0031-9228-0003-230-0, 2000, T Görnitz, D Graudenz, C.F. von Weizsäcker. *Quantum Field Theory of Binary Alternatives*, Int J Theor Phys 31, 1929-1959, https://doi.org/10.1007/BF00671965, 1992.

T Görnitz, O Ischebeck C.F. von Weizsäcker. *An Introduction to Carl Friedrich von Weizsäcker's Program for a Reconstruction of Quantum Theory*, Springer, Berlin, Heidelberg, 2003. 263-279, 2003.

T Görnitz. *Eine Entgegnung auf den Artikel 'Am Anfang war das Bit'*, Neue Züricher Zeitung, 16.02.2019, 2019.

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

H. Gould, J. Tobochnik. *Statistical and Thermal Physics: with Computer Applications*, Princeton University Press, 2010.

D. Hallyday, R. Resnick, and J. Walker. *Fundamentals of Physics, Chapter 33*, University of Toronto, 2005.

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

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.

C. Jansson. *Quantum Information Theory for Engineers: Free Climbing through Physics and Probability*, DOI: https://doi.org/10.15480/882.2285, URI: http://hdl.handle.net/11420/2781, 2019.

H. Kleinert. *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets*, World scientific, 2009.

H. Lyre. *Quantum Theory of Ur-objects as a Theory of Information*, International Journal of Theoretical Physics 34.8: 1541-1552, 1995.

H. Lyre. *Quantentheorie der Information: zur Naturphilosophie der Theorie der Ur-Alternativen und einer abstrakten Theorie der Information*, mentis Verlag, 2004.

M.A. Nielsen, and I.L. Chuang. *Quantum Computation and Quantum Information*, Cambridge University Press, 2010, `http://www.johnboccio.com/research/quantum/notes/QC10th.pdf`.

R. Penrose. *The Road to Reality: A Complete Guide to the Laws of the Universe*, Vintage, 2005.

R. Penrose. *Cycles of Time: An Extraordinary New View of the Universe*, Bodley Head, UK, 2010.

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

H. Poincaré. *Calcul des probabilitiés*, Vol. 1. Gauthier-Villars, 1912.

J. Preskill. *Physics 229: Advanced Mathematical Methods of Physics, Quantum Computation and Information*, Caltech, 1998; http://www. theory.caltech.edu/people/preskill/ph229/, 1998.

J.J. Sakurai. *Modern Quantum Mechanics*, Addison-Wesley, 1994.

C. Schiller. *Motion Mountain*, www.motionmountain.net, 2016.

D. N. Schramm. *The Big Bang and Other Explosions in Nuclear and Particle Astrophysics*, World Scientific, 1996.

D. V. Schroeder. *An Introduction to Thermal Physics*, ADDISON-WESLEY , 1999.

M. Schwartz. *Statistical Mechanics, Physics 181*, Harvard University 2018.

C. Schwarz. *Statistical Physics*, *Lecture Script, Heidelberg University*, 2017

G. Shafer, V. Vovk. *The Sources of Kolmogorov's Grundbegriffe*, Statistical Science 2006, Vol. 21, No. 1, 70-98, DOI: 10.1214/088342305000000467 c, 2006.

C. Shannon, *The Mathematical Theory of Communication*, Bell System Technical Journal, 27: 379-423, 1948.

Simon. *It from Qubit : Simons Collaboration on Quantum Fields , Gravity , and Information*, https://pdfs.semanticscholar.org, 2015

R. Swendsen. *An Introduction to Statistical Mechanics and Thermodynamics*, Oxford University Press, 2020.

M. Tribus, E.C. McIrving. *Energy and Information*, Scientific American, 225: 179- 188, 1971.

C.F. von Weizsäcker. *Komplementarität und Logik,*, Die Naturwissenschaften 42, 521-529, 545-555 , 1955.

C.F. von Weizsäcker. *Die Quantentheorie der einfachen Alternative (Komplementarität und Logik II)*, Zeitschrift für Naturforschung 13a, 245-253 , 1958.

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

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

C.F. von 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.

F. A. Wheeler, *Information, Physics, Quantum: the Search for Links*, (W. Zurek, ed.), pp. 309-336, Westview Press, 1990.

N. Wolchover. *Does Time Really Flow? New Clues Come from a Century-old Approach to Math.*, Quanta Magazine. 2020.

N. Wolchover. *Quantum Mischief Rewrites the Laws of Cause and Effect.*, Quanta Magazine. 2021.

C. Wood. *Imaginary Numbers May Be Essential for Describing Reality*, Quanta Magazine. 2021.

W.K. Wootters. *Statistical Distribution and Hilbert Space*, Phys. Rev. 23 (1981) 357-362, 1981.

A. Zee. *Magische Symmetrie: die Ästhetik in der modernen Physik*, Insel Verlag, 1993.

A. Zee. *Fearful Symmetry: The Search for Beauty in Modern Physics*, Vol. 79. Princeton university press, 2015.

A. Zeilinger. *Experiment and the Foundations of Quantum physics*, Rev. Mod. Phys. 71 (2): S288-S297 , 1999.

B. Zwiebach. *A First Course in String Theory*, Cambridge University Press, 2004.

P. Zupanovi, D. Kui. *Relation Between Boltzmann and Gibbs Entropy and Example with Multinomial Distribution*, J. Phys. Commun. 2 , 2018.

Index