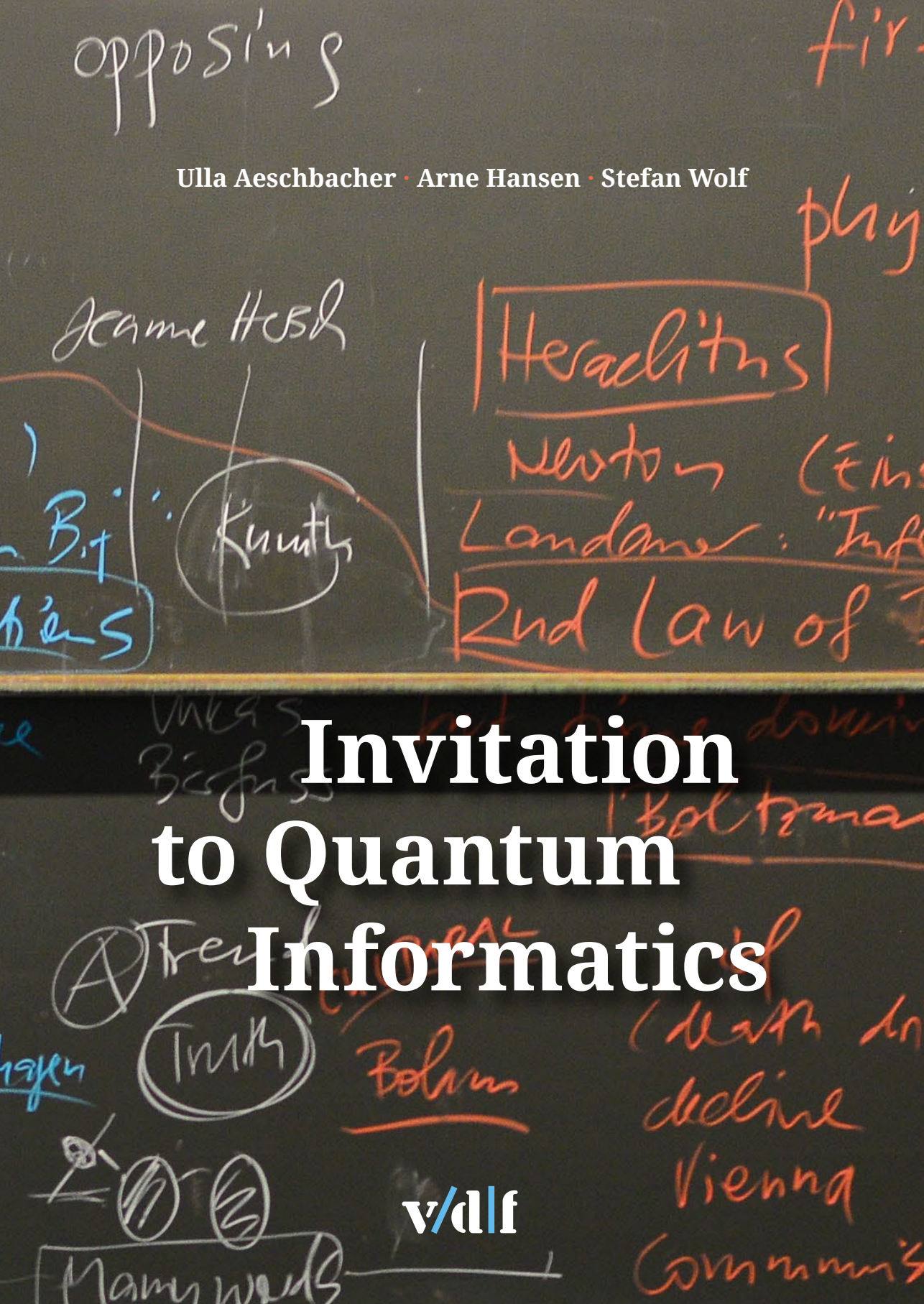
Invitation to Quantum Informatics



Ulla Aeschbacher · Arne Hansen · Stefan Wolf

Invitation to Quantum Informatics

v/dlf

Ulla Aeschbacher · Arne Hansen · Stefan Wolf

Invitation to Quantum Informatics

v/d|f

Bibliographic Information published by Die Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Internet at
http://dnb.dnb.de.

ISBN 978-3-7281-3988-7 (Printausgabe)

Download open access:
ISBN 978-3-7281-3989-4 / DOI 10.3218/3989-4

www.vdf.ethz.ch
verlag@vdf.ethz.ch

Contents

1 What Is Quantum Informatics? 5
- 1.1 Information & Physics . . . 5
- 1.2 The Stern/Gerlach Experiment . . . 7
  - 1.2.1 Independent Measurements? . . . 7
  - 1.2.2 Superposition . . . 9
- 1.3 Quantum Key Distribution . . . 10
- 1.4 The Double-Slit Experiment . . . 11
  - 1.4.1 The Mach/Zehnder Interferometer . . . 12
- 1.5 The Quantum Bit . . . 14
- 1.6 Deutsch's Algorithm . . . 15
- 1.7 The Aspect/Gisin/Zeilinger Experiments . . . 17

2 Information Is Physical 23
- 2.1 Thermodynamics and Entropy . . . 24
- 2.2 Information Theory . . . 26
  - 2.2.1 Standard Model of Communication . . . 26
  - 2.2.2 The Game of 20 Questions . . . 27
  - 2.2.3 Connection to Probability Theory . . . 28
- 2.3 The Converse of Landauer's Principle . . . 33
- 2.4 Bennett's Solution to the Problem of Maxwell's Demon . . . 34
- 2.5 Reversible Computing . . . 35
- 2.6 The Toffoli Gate . . . 38

3 Key Experiments and Postulates of Quantum Physics 43
- 3.1 Black-Body Radiation . . . 43
- 3.2 Photoelectric Effect . . . 44
- 3.3 Wave-Particle Dualism . . . 46
- 3.4 Observables . . . 48
- 3.5 Postulates of Quantum Theory . . . 50
  - 3.5.1 The State . . . 50
  - 3.5.2 The Time Evolution . . . 51

3.5.3 Observables . . . 51
3.5.4 Joint Systems and Composition . . . 52
3.5.5 Abstraction and Simplification . . . 53
3.5.6 Density Matrices . . . 55
3.6 Qbits . . . 58
3.6.1 One Qbit . . . 58
3.6.2 Two Qbits . . . 59
3.6.3 The CNOT Gate . . . 64
3.6.4 Cloning, Pseudo-Cloning, and Pseudo-Measurements . . . 67

4 Quantum Communication **71**
4.1 Teleportation . . . 71
4.2 Superdense Coding . . . 74

5 Simple Algorithms **75**
5.1 n Qbits . . . 75
5.2 The Secret Mask . . . 76
5.3 The Deutsch/Josza Algorithm . . . 79

6 Pseudo-Telepathy **83**

7 The Needle in the Haystack: Grover's Algorithm **87**
7.1 Motivation . . . 87
7.2 The Elements . . . 88
7.3 The Grover Circuit . . . 88

8 Integer Factoring: Shor's Algorithm **91**
8.1 Quantum Fourier Transform . . . 91
8.2 Phase Estimation . . . 93
8.3 Factoring . . . 94

9 Epilogue: Information & Physics **97**

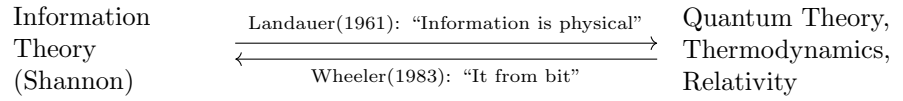
10 Bibliography **117**

Chapter 1

What Is Quantum Informatics?

1.1 Information & Physics

Physics and *Information (Theory)* are two different sciences, *i.e.*, two thinking traditions both rooted in their respective histories, coined by their own methods, personalities, and established truths. The present text belongs to the (postmodern) tradition of considering, establishing, discussing, and analyzing the connections between physics and information. Of these connections, there are essentially two natures: On the one hand, experience, observation, and physical discourse are in the form of information: *John Archibald Wheeler* compressed this fact to the slogan "It from Bit." On the other hand, information representation, processing, and transmission are, ultimately, *physical* processes; as *Rolf Landauer* put it: "Information is Physical."



This text starts from the latter insight and discusses consequences thereof both of limiting (thermodynamics) and enabling (quantum theory) character of physical law for information treatment. On occasion, a glimpse is offered at the possibility of obtaining new insights into natural law when the informational point of view is chosen. The text culminates in *Peter Shor*'s algorithm, born out of a surprising and breathtaking marriage between quantum physics and number theory. (*Claus Hepp* called the algorithm the "most fascinating result in theoretical physics of its decade," due to its internal conceptual beauty,

not its "real-world" application that is, at this point, potential, unclear, and debated.)

Concretely, Landauer's slogan means that the representation of a bit of information, if this bit is to "exist," must be physical. This implementation can be realized by a switch, a current in a metal wire going one way as opposed to the other, the position of a single gas molecule in a container, the polarization of a photon, or by the electron of a hydrogen atom in its ground as opposed to first excited state. The latter example is interesting since it illustrates that *digitalization* in fact comes very naturally with quantization (*e.g.*, of energy levels) whereas in classical physics, it has to be enforced in some way. Later in the text, however, we will see that quantum physics allows for another kind of "world between zero and one" that could more accurately be enabled by the possibility of "being zero and one at the same time (at least to some extent)."

If we follow that thought through, we realize that physical laws thus can have direct consequences for information processing. Although that is true in principle, it seems that the nature of these consequences probably depends strongly on the specific choice of the information's physical representation. Or — to turn that thought around — are those physical laws that have consequences that are *independent* of that representation (beyond the fact that there *is* such a representation) perhaps laws that are rather logical-informational than "physical" in the strict sense?

The second law of thermodynamics states that, in a closed system, *entropy* does not decrease (with overwhelming probability). What is entropy? A first, rough answer is that it is some kind of measure for *disorder*. A precise answer is harder; *John von Neumann* was quoted as saying "if you want to win any discussion, just say 'entropy' and you will be on the safe side, because nobody really knows what entropy is." It has also been said that *Claude Shannon*, the founder of information theory, followed von Neumann's advice when he chose the name "entropy" for the central quantity of his theory.

A remarkable feature of the second law is its *time asymmetry*, which contrasts the time symmetry of most physical laws and processes. Exceptions are some elementary-particle reactions and, more importantly for us, *measurements*. Related notions thus would be *past* and *future*: the arrow of time.

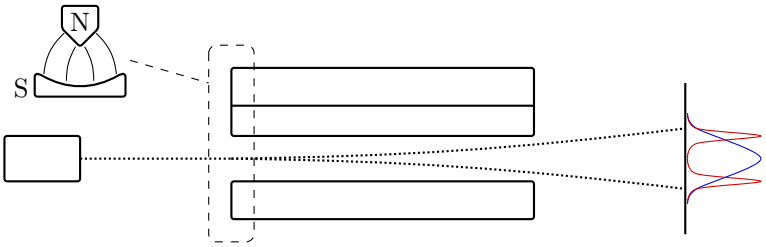
Whereas entropy (disorder) may be hard to define in general, it is clear in some cases: Given that N binary memory cells contain a "random" content (an equally problematic notion, in fact) and are then all erased (put to 0), the entropy in the set of memory cells drops.

1.2 The Stern/Gerlach Experiment

1.2.1 Independent Measurements?

The Stern/Gerlach experiment — proposed by Otto Stern in 1921 [16] and carried out by Walther Gerlach in 1922 [8] — was not the first in the history of quantum theory, but one of the most important ones to understand the structure and properties of the basic building block of quantum information processing, the *quantum bit (Qbit)*. In particular, the question was what *classical* information we can get on such a Qbit, and how.

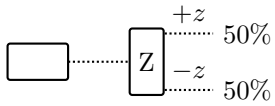
In the experiment, Stern and Gerlach measure a certain quantity, the *magnetic dipole moment*, of silver atoms by sending a stream of such atoms, exiting an oven, through a inhomogeneous magnetic field. Each atom is then deflected from the path proportionally to its dipole in the direction of the magnets. If we imagine that the moments of the atoms point in random directions (and have, perhaps, constant length or even varying length within some range or according to some distribution), then the classical expectation is that the deflection pattern reaches a maximum in the middle (no deflection) and then symmetrically, monotonically, and continuously decreases on the sides. This is, however, not what was observed: There is no detection in (not even close to) the middle, but rather two sharp peaks at equal distances from that middle.



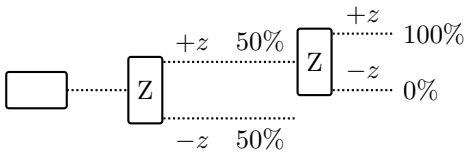
This "quantization" is one of the characteristic features of quantum theory — to which it owes its name, too — and motivated assigning the quantity a new name in that context: *spin*.[1]

[1]The following anecdote was reported concerning this experiment: Initially, Gerlach did not see any detection of the screen supposed to register the trajectories of the silver atoms. Desperately, he handed the blind plates to Stern, who gave it a look to; during that, some of the air Stern was breathing out hit the plates. The thing is that the cigars Stern used to smoke (heavily) contained a lot of sulfur; they were cheap cigars, as physics researchers were not well paid at the time, it seems. In the end, the sulfur initiated the reaction necessary to see the detections on the screen, and the experiment succeeded. The story is sometimes taken to support the argument that also social and economic factors (Stern's salary and the quality of his cigars, etc.) have to be considered in the context of physical experiments dismantling "objective" reality.

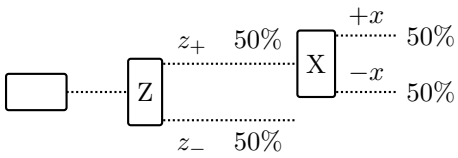
In the case of a single Stern/Gerlach measurement, say, in the Z-direction, two identical rays result. Let us call the rays by the properties they correspond to, i.e., $Z-$ and $Z+$.



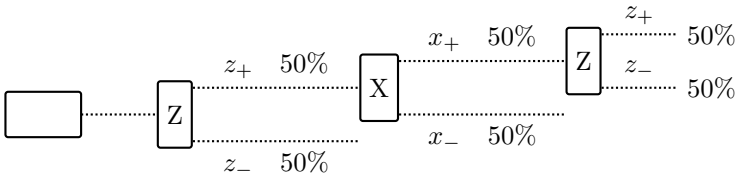
If the same measurement is repeated on, say, only the $Z+$ ray, then all atoms are again deflected in the $+$ direction. In this sense, the Z-spin property looks classical: It is stable with respect to repeated measurements.



When the magnet is rotated into the spatial X direction (also perpendicular to the flying direction Y of the atoms), then a $50-50$ distribution arises. This is not surprising due to the geometrical symmetry of the situation. It is equally unsurprising that the same is observed when the second measurement (the one in X direction) is carried out *after* a Z measurement from which only the $Z+$ counts are carried over to the next experiments: It means that the two properties, "Z-spin" and "X-spin," look *independent*.



The most fascinating outcome results when the two types of measurement are combined as follows: First, a Z measurement, whereby only $Z+$ counts are transferred to the next magnet, an X measurement. If subsequently, another Z measurement is performed, then half the particles show $Z-$ spin, although we took only $Z+$ states after the first measurement. This is puzzling and questions both our interpretations above: The *stability* as well as the *independence* of the properties in question.



Interlude

The *stability of a measurement result* is not so surprising: Popper regards scientifically interesting physical effects to be defined by being reproducible by anyone and at anytime, provided that one builds the same experimental setup.[a] The "scientific method" crucially relies on being able to enquire about equivalent questions and then expect the same answer. There must at least exist some conditions under which this is possible. This does, however, not imply that this is possible under all conditions as one might hope coming from classical mechanics.

[a] "Der wissenschaftlich belangvolle physikalische Effekt kann ja geradezu dadurch definiert werden, daß er sich regelmäßig und von jedem reproduzieren läßt, der die Versuchsanordnung nach Vorschrift aufbaut." [14, §I.8]

1.2.2 Superposition

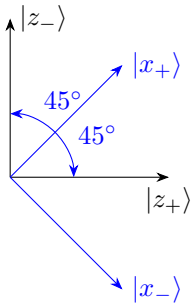
The statistics found within the Stern/Gerlach experiment were surprising for *single* particles. They would not have been surprising if we had dealt with waves. Imagine a polarizing filter in a beam of light. Then measuring z_+ can be considered to correspond to passing a polarizing filter of a certain orientation; measuring z_- corresponds to passing a polarizing filter rotated by 90°. If the beam initially is unpolarized, then the probability of passing such a polarizing filter—or the ratio of the intensity before and after the filter—is 50%. Measuring with a filter rotated by 45° with respect to the z_+ filter would correspond to x_+. Then the intensities measured in a sequence of filter would fit the probabilities in the Stern/Gerlach experiment.

The essential property of waves is that they can be linearly combined. Quantum mechanical states have the same property: They are elements in a vector space. But their length does not relate to wave amplitudes; instead, they serve to derive probability distributions. Thus, linearly combined states have to be normalized. The z_+ filter then corresponds to asking whether the silver atom is in a z_+ state, denoted by $|z_+\rangle$. If the silver atom does not go up, i.e., it is not in a state $|z_+\rangle$, then it goes down, i.e., it is in a state $|z_-\rangle$, orthogonal to $|z_-\rangle$. So, the question whether the silver atom is in the state $|z_+\rangle$ and the question whether the silver atom is in the state $|z_-\rangle$ are complementary to one another. In fact, they can also be regarded as two different answers to the same question, i.e., the Z measurement.

If, after a Z measurement, we perform an X measurement, then we want to know whether the silver atom is in a state $|x_+\rangle$ or in a state $|x_-\rangle$. Both are equal superpositions,

$$|x_+\rangle = \frac{1}{\sqrt{2}}|z_+\rangle + \frac{1}{\sqrt{2}}|z_-\rangle \qquad |x_-\rangle = \frac{1}{\sqrt{2}}|z_+\rangle - \frac{1}{\sqrt{2}}|z_-\rangle.$$

No matter whether we had obtained z_+ or z_- in the Z measurement, the X measurement yields one of both results with equal probability.[2] Also in the inverse order: A Z measurement after an X measurement yields the same uniform distribution—independent of any measurements before the X measurement.



Interlude

So, a phenomenological perspective, i.e., from a comparison of probability distributions, suggests the superposition of states in quantum mechanics. Quantum mechanics attains an essential property of wave mechanics, even though there are no more coupled system, with a description in, e.g., classical mechanics. The states are then more abstracts entities. They are no longer directly observable properties of a system, but rather tools to determine probability distributions for measurement results.[a]

[a] *Grete Hermann* describes quantum states as "new symbols that express the mutual dependency of the determinability of different measurements." [10]

1.3 Quantum Key Distribution

Previously we have seen: The condition for measuring *with certainty* the same value in two consecutive measurements with the same measurement basis, e.g., in two consecutive Z measurements, is that there is no intermediary measurement in another bases. In other words: The interactions of a system with its environment, within, say, a measurement, become traceable. This allows us to detect an eavesdropper in a cryptographic key agreement protocol. In 1984, Gilles Brassard and Charles Bennett developed the first application of quantum mechanics for cryptographic purposes with such a key agreement protocol [2].

Let us assume that Eve and Bob can exchange quantum mechanical systems. Then they can establish a secret key as follows: Alice chooses at random

[2] The details of how to derive probabilities from states will be given later.

a measurement, either Z or X, and measures a quantum system, e.g., a silver atom, in that basis. She then sends that system to Bob, who also chooses at random between a Z and an X measurement, and performs the measurement on that system. If the bases that Alice and Bob choose coïncide, then the results of their measurements are the same—unless there has been an eavesdropper, Eve, measuring the system during its transmission from Alice to Bob in a basis different from Alice's and Bob's. Alice and Bob do *not* agree beforehand on a basis. Instead, they repeatedly measure quantum systems in randomly chosen bases. So, Eve can merely guess Alice's choice of measurement. If Alice's choice was really random then, in some cases, Eve guesses wrongly and, therefore, disturbs the system. Alice and Bob can trace that disturbance as follows: Alice repeatedly chooses random measurement and sends the states after the measurement over to Bob, e.g.,

Alice's measurement	×	+	+	×	+	×	×	+	×
result	0	0	1	0	1	1	0	0	1

.

Bob also chooses his bases at random and measures the state:

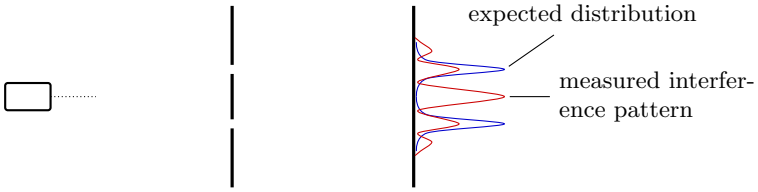
Alice's measurement	×	+	+	×	+	×	×	+	×
result	0	**0**	**1**	0	1	**1**	0	**0**	1
Bob's measurement	+	+	+	+	×	×	+	+	+
result	1	**0**	**1**	0	0	**1**	1	**0**	1

.

Where their bases agree, their measurement result are the same, if there is no eavesdropper. So, Alice and Bob communicate over an authenticated channel the positions in the above sequence where they do agree. Now, to ensure that there has been no eavesdropper, they finally choose randomly some of the positions where their results should be the same and compare whether they actually are. If Eve had been intercepting and measuring the states, then the results should differ in about 1/4 of the cases. If Alice and Bob find that their results are the same in (almost) all cases, then they can use the remaining, unpublished measurement result (where their measurement bases agree) as a secret key.

1.4 The Double-Slit Experiment

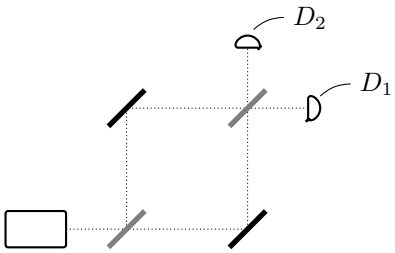
If one shines light onto a double slit, an *interference pattern* appears on a screen behind the double slit. What happens, however, if one sends *single* electrons or *single* photons onto the double slit? Intuitively one would expect two peaks, corresponding to each of the slits. Instead, if one measures the position of the electrons or photons on the screen for many repetitions of the experiment, an interference pattern emerges.



Surprisingly, single particles exhibit wave properties. In a sense, the different paths *the particle could have taken* interfere with one another. If one measures, however, which path the particle has taken, then the interference pattern vanishes.

1.4.1 The Mach/Zehnder Interferometer

The Mach/Zehnder interferometer can be considered a variant of the double-slit experiment. If one sends single photons into a Mach-Zehnder interferometer



then interference occurs, and the photon is detected with certainty in detector D_1. More precisely: In each reflection, the state of the photon picks up a phase shift of $\pi/2$. If we label the state of a photon moving to the right by $|1\rangle$ and a photon moving up by $|2\rangle$, then the effect of the fully-reflecting mirrors is

$$|1\rangle \mapsto i|2\rangle \qquad |2\rangle \mapsto i|1\rangle\,,$$

and the effect of the semitransparent mirrors is

$$|1\rangle \mapsto \frac{1}{\sqrt{2}}\left(|1\rangle + i|2\rangle\right) \qquad |2\rangle \mapsto \frac{1}{\sqrt{2}}\left(|2\rangle + i|2\rangle\right)\ .$$

This characterizes two linear maps that allow tracing the state of photon as it moves through the interferometer after its emission from the source

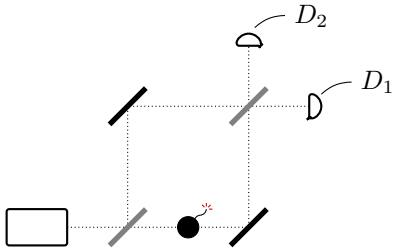
$$\begin{aligned} |1\rangle &\mapsto \frac{1}{\sqrt{2}}\left(|1\rangle + i|2\rangle\right) \mapsto \frac{1}{\sqrt{2}}\left(i|2\rangle - |1\rangle\right) \\ &\mapsto \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\left(i|2\rangle - |1\rangle\right) - \frac{1}{\sqrt{2}}\left(|1\rangle + i|2\rangle\right)\right) = -|1\rangle\ . \end{aligned}$$

With a probability of $1/2$, the photon passes the semitransparent mirror without being affected, and with the same probability it is reflected and picks up a phase shift of $\pi/2$. So after the second semitransparent mirror, the photon is in a state $-|1\rangle$, and will be measured with certainty in detector D_1.[3]

If, however, one measures whether the photon has gone through the upper or the lower arm of the interferometer, then the state before the last semitransparent mirror is *either* $|1\rangle$ *or* $|2\rangle$. In that semitransparent mirror the state of the photon is then mapped to either $1/\sqrt{2}(|1\rangle + i|2\rangle)$ or $1/\sqrt{2}(|2\rangle + i|1\rangle)$. In both cases the photon is detected in either of the detectors with equal probability. Thus, a measurement about the path of the photon affects the interference. This effect can be used to detect explosive bombs.

Interlude: Interaction-free measurements

In 1993, Avshalom Elitzur and Lev Vaidman [7] proposed a method for *measuring without interacting*, employing a Mach/Zehnder interferometer in the following way. Imagine a bomb that is triggered by a single photon. How could one detect such a bomb? Looking at it would expose it to photons and thus explode it. There is a way around it. *Literally.* If one puts the bomb into one arm of a Mach-Zehnder interferometer, then one can detect a photon if it travels through the *other* arm.



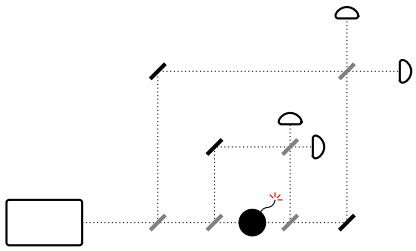
The bomb measures which path the photon has taken and thus affects the interference: The photon can now be measured in the detector D_2 with probability $1/4$, where, without the bomb, it would not have been measured at all. If, in the first measurement, the photon was detected in D_1, we can simply send another photon. Then the probability to detect the photon in D_2 without exploding the bomb is $1/4 + 1/4^2$. Proceeding

[3] The minus sign is irrelevant for the probability distribution, as we will see later in the course.

this way, we can reach a probability of

$$\sum_{n=1}^{\infty} 1/4^n = 1/4 \sum_{n=0}^{\infty} 1/4^n = 1/3.$$

But the bomb still explodes with probability 1/2. How could we reduce this threat? For instance, we could encapsulate the multiple interferometer:



Another equivalent way of reducing the threat of explosion is to make the first semitransparent mirror almost *in*transparent and the second almost transparent.

1.5 The Quantum Bit

To transfer a bit $b \in \{0, 1\}$ into the quantum world, we associate 0 and 1 with two orthogonal vectors, usually with the standard basis vectors,

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

A general quantum state can then be written as a superposition $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, with $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. Measuring $|\psi\rangle$ in the standard basis yields 0 with probability $|\alpha|^2$ and 1 with probability $|\beta|^2$. Quantum circuits are composed of quantum gates, i.e., unitary maps. The most important of these is the Hadamard gate,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

which maps the basis states to superpositions,

$$H|0\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \qquad H|1\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \ .$$

Applying the Hadamard gate again, yields the standard basis vectors.

Another interesting gate is

$$F = \frac{1}{\sqrt{2i}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

because applying the gates twice, yields the not-gate,

$$F \cdot F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

Classically there is no gate that yields the not-gate this way: F has been called the "square root of NOT,"

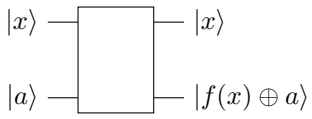
$$F = \sqrt{\text{NOT}} .$$

1.6 Deutsch's Algorithm

Interference is an essential ingredient in Deutsch's algorithm. Let us assume we are given a black box that implements a function

$$f : \{0,1\} \to \{0,1\} .$$

The question is: Is f constant or not? Is $f(0) \oplus f(1)$ zero or one? Classically, we have to query the black box twice to find out. But we can do better with the help of quantum mechanics. To achieve that we first have to turn the black box into a quantum black box. Because of the unitarity of the quantum mechanical time evolution, the quantum version has to be reversible. Thus, we assume that the quantum black box implementing f is a gate of the form



So if a is 0, then x is mapped to $|f(x)\rangle$ on the output wire. And if a is 1, then x is mapped to $|\overline{f(x)}\rangle$, where the overline indicates the negation. The box characterizes a map by linear extension. A first idea might be to put a superposition on the input wire and set a to 0. We have to consider both input wires together[4] and obtain the combined input

$$\begin{aligned} &\frac{1}{\sqrt{2}} \left(|0\rangle + |1\rangle\right)) \otimes |0\rangle \\ = &\frac{1}{\sqrt{2}} \left(|0\rangle \otimes |0\rangle + |1\rangle \otimes |0\rangle\right)) . \end{aligned}$$

[4]See Section 3.5 for tensor products describing combined systems.

The first summand is mapped to $|0\rangle \oplus |f(0)\rangle$ and the second to $|1\rangle \oplus |f(1)\rangle$. Linearly combining the two we obtain

$$\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)) \otimes |0\rangle \;\mapsto\; \frac{1}{\sqrt{2}}\left(|0\rangle \oplus |f(0)\rangle + |1\rangle \oplus |f(1)\rangle\right)) \;.$$

The resulting state is entangled, and we cannot access the information about $f(0)$ and $f(1)$ by merely measuring the output wire.

The magic happens when we put a superposition on the second wire as well. If we set $|a\rangle = 1/\sqrt{2}\,(|0\rangle - |1\rangle)$, then we can expand the combined input as

$$\begin{aligned}&\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right)\\ =\;&\frac{1}{2}\big(|0\rangle \otimes |0\rangle - |0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle - |1\rangle \otimes |1\rangle\big)\,.\end{aligned}$$

Applying the gate to each of the summands and linearly combining the result yields

$$\begin{aligned}&\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)) \otimes \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right))\\ \mapsto\quad&\frac{1}{2}\left(|0\rangle \otimes |f(0)\rangle - |0\rangle \otimes |\overline{f(0)}\rangle + |1\rangle \otimes |f(1)\rangle - |1\rangle \otimes |\overline{f(1)}\rangle\right)\\ =\;&\frac{1}{2}\left(|0\rangle \otimes \left(|f(0)\rangle - |\overline{f(0)}\rangle\right) + |1\rangle \otimes \left(|f(1)\rangle - |\overline{f(1)}\rangle\right)\right)\end{aligned}$$

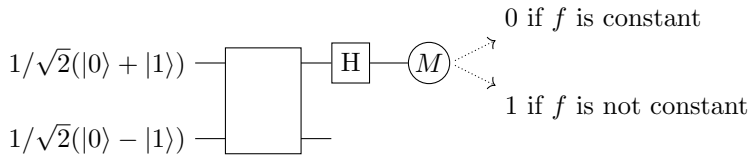
So, if $f(0) = f(1)$, then the last line is equal to

$$\frac{1}{2}\left(|0\rangle + |1\rangle\right) \otimes \left(|f(0)\rangle - |\overline{f(0)}\rangle\right)\;;$$

otherwise

$$\pm\frac{1}{2}\left(|0\rangle - |1\rangle\right) \otimes \left(|f(0)\rangle - |\overline{f(0)}\rangle\right)\;.$$

Then, measuring in the standard basis after applying a Hadamard gate yields 0 if the first is the case and 1 if the latter is the case.

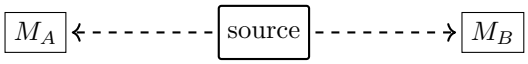


Deutsch's algorithm does not allow us to retrieve *more* information. The last measurement still just returns one bit. It yields the actual value of neither $f(0)$ nor of $f(1)$. Rather, the algorithm allows accessing the *right* bit, i.e., $f(0) \oplus f(1)$.

1.7 The Aspect/Gisin/Zeilinger Experiments

The power of quantum computing and quantum information processing lies in the fact that *a pair of Qbits* is more than "one Qbit plus another Qbit." A particularly striking manifestation of that additional quality is revealed in the correlations that arise when two Qbits which are in a so-called *entangled state* are independently measured.

Imagine that, in a preparation central, pairs of Qbits (*e.g.*, polarized photons) are generated and sent onto their respective paths to two parties, Alice and Bob. They measure the particles, for instance, both in the standard basis, and observe (when they compare their data) certain correlations, for instance, the same bit in every run.



Assume for the sake of the argument that this is exactly what they see: The same bit in every run. Is this surprising? Is there something mysterious about it? *A priori* not at all: We are surrounded by correlations, and we are often interested in how they arose. Let us give it a try in the given situation: One possibility is that, in the preparation center, a coin in flipped in each run, and according to the outcome either a QBit in state $|0\rangle$ is sent to both parties (we call this state $|0\rangle \otimes |0\rangle$, where for the moment, the symbol $\otimes$ is simply to be read as "and:" Alice receives $|0\rangle$ *and* Bob receives $|0\rangle$), or $|1\rangle$ to both (*i.e.*, state $|1\rangle \otimes |1\rangle$). When they then measure the respective particles, they have their perfect correlation. But this is not how it was done in these experiments; it would also not have been very interesting, by the way.

What *was* actually sent and measured are the two parts of the *equal superposition* of the above two states and not their "classical-probabilistic" mixture:

$$\frac{|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle}{\sqrt{2}} \ .$$

Since the probability amplitudes of the two basis vectors corresponding to *equal* outputs of Alice and Bob are both $1/\sqrt{2}$, whereas those of the joint states $|0\rangle \otimes |1\rangle$ as well as $|1\rangle \otimes |0\rangle$ are zero, the outcome is the same: If Alice and Bob both measure in the "standard" basis $\{|0\rangle, |1\rangle\}$, then they always receive the same output: a perfect correlation.

Is this now mysterious and strange? According to quantum theory, the outputs are measurements are *truly random*. When combined with that, the correlation *does* indeed look strange: How can the outcomes of Alice and Bob be perfectly correlated and at the same time spontaneously arise upon measurement? The weirdness of this idea, and the desire to explain also this correlation in the traditional ways, motivated *Boris Podolsky* and two co-authors

to write that quantum theory was incomplete [6]. In fact, if we imagine that the future measurement outcomes are already determined in the preparation centre, and then sent along the particles — encoded into them somehow in a "hidden" way: *hidden variables* — then the mystery around the correlations immediately disappears.

Claim (Einstein, Podolsky, and Rosen, 1935)**.** *Quantum theory is* incomplete *and must be augmented by "hidden parameters" completely determining the measurement outcomes of all alternative measurements.*

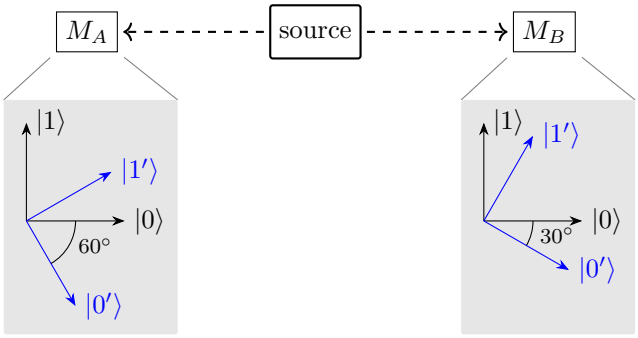
It is ironic that the *exact same states, and the correlations they give rise to*, were recognized almost 30 years later to suggest a striking argument *against* hidden-variable preparations of measurement results. How can that be? The correlations in question — which were the motivation in the first place for "EPR" to ask for pre-distributed pieces of classical information —, these exact same correlations are in fact *too strong* to be fully explained by such a mechanism. This insight from *John Stewart Bell* in 1964 [1] resulted when Bell gave Alice and Bob more liberty and had them not always do a measurement in the standard basis $\{|0\rangle, |1\rangle\}$ but gave them a choice to use another orthogonal basis instead. (Remember that our lesson from the Stern/Gerlach experiment was, whereas we can *choose* between one of two measurements, it could *not* be expected that the result of the second would still have anything to do with what happens in the preparation center, or what is in the other party's hands when both are carried out, one after the other.)

Claim (Bell, 1964)**.** *EPR's program is* in doubt*: There exist quantum correlations that go beyond the explanatory power of shared classical information.*

In order to understand Bell's reasoning leading up to that claim, we have to investigate the joint measurement statistics of maximally entangled states under different choices of measurement basis by Alice and Bob. We look here at a different but related state to the one above, namely, the *"singlet"*:

$$|\Psi^-\rangle := \frac{|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle}{\sqrt{2}} \ .$$

This state has, compared to other "maximally entangled states," nicer transformation properties: For instance, when Alice and Bob choose *the same general* basis, then what statistics do they observe? In order to figure this out, we rewrite the singlet in a basis $\{|0'\rangle, |1'\rangle\}$ rotated with respect to the standard basis by some angle α.



The linear basis-transformation map is then a rotation:

$$\begin{aligned}|0\rangle &= \cos(\alpha)|0'\rangle - \sin(\alpha)|1'\rangle \\ |1\rangle &= \sin(\alpha)|0'\rangle + \cos(\alpha)|1'\rangle \ .\end{aligned}$$

We denote here the vectors of the standard basis in terms of the rotated basis; this way, we can simply replace $|0\rangle$ and $|1\rangle$ in the singlet's definition (let $c := \cos(\alpha)$, and $s := \sin(\alpha)$):

$$\begin{aligned}|\Psi^-\rangle &= \frac{1}{\sqrt{2}}\Big((c|0'\rangle - s|1'\rangle)\otimes(s|0'\rangle + c|1'\rangle) - (s|0'\rangle + c|1'\rangle)\otimes(c|0'\rangle - s|1'\rangle)\Big) \\ &= \frac{1}{\sqrt{2}}\Big(|0'\rangle\otimes|0'\rangle(cs - sc) + |0'\rangle\otimes|1'\rangle(c^2 + s^s) \\ &\qquad + |1'\rangle\otimes\ |0'\rangle(-s^2 - c^2) + |1'\rangle\otimes|1'\rangle(-sc + cs)\Big) \\ &= \frac{1}{\sqrt{2}}\Big(|0'\rangle\otimes|1'\rangle - |1'\rangle\otimes|0'\rangle\Big) \ :\end{aligned}$$

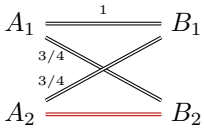
The singlet written with respect to a general basis has the same form as in the standard basis. In particular: If both parties measure in the same basis, they get a perfect anti-correlation in their results.

What happens if Alice and Bob *do not* measure in the same basis? In order to see this, assume now that only Bob rotates his basis by α, whereas Alice uses the standard basis. The state can again be rewritten in the respective measurement basis:

$$\begin{aligned}|\Psi^-\rangle &= \frac{1}{\sqrt{2}}\Big(|0\rangle\otimes(s|0'\rangle + c|1'\rangle) - |1\rangle\otimes(c|0'\rangle - s|1'\rangle)\Big) \\ &= \frac{1}{\sqrt{2}}\Big(s|0\rangle\otimes|0'\rangle + c|0\rangle\otimes|1'\rangle - c|1\rangle\otimes|0'\rangle + s|1\rangle\otimes|1'\rangle\Big) \ .\end{aligned}$$

The probability now that Alice and Bob receive opposite output bits is $\cos^2(\alpha)$.

We have now analyzed the state closely enough to be able to test EPR's claim that all outputs of the alternative measurements are predetermined. For this, we consider the scenario where Alice chooses between the standard basis and the basis rotated by $\alpha = 30°$. We assume further that Bob can choose between the standard basis and the basis rotated by $\alpha = -30°$, *i.e.*, in the other sense. For simplicity's sake, we also assume that Bob flips his output bit: In the case where both measure in the standard basis, we then have a perfect *correlation*, not an *anti*correlation. Let us now assume, in the spirit of *Einstein et al.*'s intervention, that the corresponding output bits A_1, A_2, B_1, and B_2 are chosen already in the preparation center. Here, the respective first bits result when the parties choose the standard basis. Does the quadruple of bits even exist? What statistics do the bits have to satisfy?



First, we must have $A_1 = B_1$ with certainty. Second, $A_1 = B_2$ must hold with the probability of exactly $\cos^2(30°) = 3/4$; the probability for $A_2 = B_1$ is the same: The bases enclose an angle of 30°. Before we go on: What can we conclude from these three facts about the probability of $A_2 = B_2$? According to the transitivity of equality and the union bound, we must have

$$\begin{aligned}\mathrm{Prob}[A_2 \neq B_2] &\leq \mathrm{Prob}[A_1 \neq B_1] + \mathrm{Prob}[A_1 \neq B_2] + \mathrm{Prob}[A_2 \neq B_1] \\ &= 0 + 1/4 + 1/4 = 1/2 \ .\end{aligned}$$

The inequality — which is the simplest example of a so-called *Bell inequality* — follows from the fact that the *inequality* of A_2 and B_2 can occur only if also (at least) one of the other three equalities fails to hold.

But what is the probability $\mathrm{Prob}[A_2 \neq B_2]$ according to quantum theory? The angle between the bases being 60°, we have

$$\mathrm{P}[A_2 = B_2] = \cos^2(60°) = 1/4 \ .$$

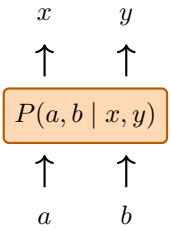
Bell's inequality is *violated*![5]

The measurement results are correlated, but nevertheless arise upon measurement and not prior to it. This deeply disturbing fact has been called *(Bell)*

[5] We are well aware of the rule "do not shout at people." However, we *must* stress this point, in the face of the strength of the fact. The fact has been experimentally tested and claimed to have been verified in various experiments, under different conditions. The correlations even appeared in a relativistic experiment where the *each particle was measured before the other in its respective basis.*

non-locality. It deeply questions basic notions we are used to, such as "space-time causality." Indeed, *Reichenbach's principle*, stating that any correlation between two events that we can observe is established either by a common cause in the common past of the event, or a direct influence from one to the other, seems incompatible with quantum correlations. Whereas Bell's result rules out a piece of classical information generated in the common past as the cause of the correlation, signalling mechanisms are, to say the least, unsatisfactory; for instance, since the speed of that influence would have to be highly superluminal, in sharp conflict with the spirit of relativity theory.

After the shock caused by this mystery, further questions came up: Do the weird correlations have applications? How strong can non-local correlations get? In order to study the phenomenon, *Popescu and Rohrlich* proposed a simple *idealized* "non-local behavior" and named it *PR-box*: It has a bipartite input-output behavior,



for which the inputs and the outputs satisfy

$$a \oplus b = x \cdot y\,.$$

So we obtain four equations,

$$\begin{aligned} a_1 \oplus b_1 &= x_1 \cdot y_1 \\ a_1 \oplus b_2 &= x_1 \cdot y_2 \\ a_2 \oplus b_1 &= x_2 \cdot y_1 \\ a_2 \oplus b_2 &= x_2 \cdot y_2 \end{aligned}$$

If we add (modulo 2) all these equations, i.e., apply the "xor" (exclusive or) to all of them together, then the left-hand side yields zero, as $a \oplus a = 0$, whereas the right-hand side yields

$$(x_1 \oplus x_2) \cdot (y_1 \oplus y_2) = 1\,.$$

So we cannot solve this system of four equations as merely three out of the four can be satisfied. Therefore, we can classically approximate this by a maximum of 75%. One assignment of values that satisfies three out of four equations is

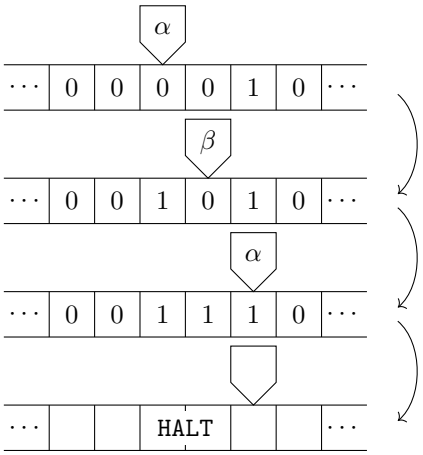
$$a_1 = 0 \quad b_1 = 1 \quad x_1 = 1 \quad y_1 = 1\,.$$

Chapter 2

"Information Is Physical" (Landauer, 1961)

If we look at a computing device purely from the point of view of physics, it essentially transforms (electrical) free energy into heat. Of course, a computer is not an oven, and this heat dissipation is not what we are interested in; the latter is harder to define purely physically.[1]

Let us consider the computing model that is standard in theoretical computer science, the *Turing machine*.

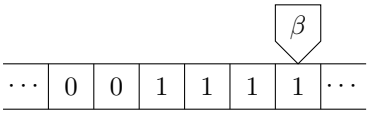


$$\begin{aligned}
\alpha\,0 &\mapsto 1\,R\,\beta \\
\alpha\,1 &\mapsto \texttt{HALT} \\
\beta\,0 &\mapsto 1\,R\,\alpha \\
\beta\,1 &\mapsto 0\,L\,\alpha
\end{aligned}$$

Let us now ask the question about that computation that is, actually, the most relevant one from the point of view of physics, as we will see later: Is the

[1] This characterization resembles Rënyi's quote: "A mathematician is a device for turning coffee into theorems."

computation *logically reversible*, *i.e.*, are the predecessor states (tape content plus position and state of the head) also uniquely determined by their successor states, just as the latter are by the former? The answer in general is, even for deterministic Turing machines, *no*.



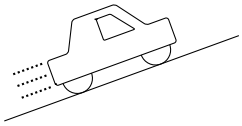
Why is this fact physically relevant? It is maybe not so much for Turing machines, but certainly for "real-world" computers. For those, we have the following "law."

Moore's Law: The performance of computers doubles every 18 months.

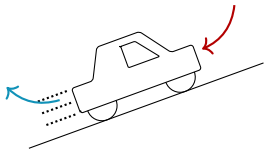
Two remarks on that: This is not exactly a physical law in the narrow sense; but then, what kind of law is it? It may also have aspects of a *self-fulfilling prophecy*. Second, like all exponential laws, it has an end. That, when, and how this end is reached has a lot to do with the laws of physics, more precisely and first of all: *Thermodynamics*; in particular its uncomfortable second law.

2.1 Thermodynamics and Entropy

Law 1 (First law of thermodynamics: Energy)**.** *A* perpetuum mobile *of the first kind is impossible. In a closed system, the total energy is* constant.



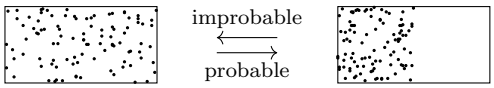
Law 2 (Second law of thermodynamics: Entropy)**.** *A* perpetuum mobile *of the second kind is also impossible. In a closed system, entropy ("disorder") does not decrease.*



A remarkable feature of the second law is that it is *asymmetric* in time, whereas, for instance, the laws of classical mechanics are not. How can a time-asymmetric law follow from time-symmetric axioms such as Newton's laws?

This is a profound question for which there does not seem to be a simple answer; it is rather a quite fascinating minefield, as it seems. What is for sure is that, in our everyday lives, time asymmetry is extremely common and normal: We do not have any problem distinguishing yesterday from tomorrow. *But why?* Another example: If we need money and smash our *piggy bank*, then it is not hard to know which is its state *before* and which is *after*.

Simple to analyze and equally constructive is the example of a *gas container* holding a (large) number of molecules.



What is interesting about the picture is that what happens when the gas transits from the compressed to the relaxed state are simply elastic collisions between molecules — each of which is by itself perfectly reversible. How can a large number of perfectly reversible microscopic processes lead to a totally asymmetric global behavior? Crucial is the "large number." This means that, in the end, the second law is rather more statistical in nature than physical in the strict sense. Does it also mean that "yesterday" and "tomorrow" are *statistical notions rather than physical ones*?

Even finding a precise formulation of the second law is already a mine field. For instance, *Poincaré's recurrence theorem* implies that a closed system starts in very ordered state and gets disordered, and that it returns to an arbitrarily close approximation of that order. However, you have to wait for this to happen for a very long time.

Let us turn the wheel of history back here and look at the history of the law. The second law of thermodynamics was discovered by *Sadi Carnot* and introduced in a text entitled "Réflexions sur la Puissance Motrice du Feu," when Carnot was 24 years old. This was his sole publication, so his H-index is 1, which is quite telling (about the index). Later versions of the second law were due to Clausius and Thomson [later Lord Kelvin]. Clausius' version implies that temperature differences tend to disappear. Furthermore, he predicted, as a consequence of his law, the "heat death" of the universe. Clausius was criticized by his pupil Max Planck, who held the view of the entire universe as a closed system to be untenable; Clausius removed the corresponding remarks from his publications. Kelvin realized a strangeness of the second law, by saying that radiation alone could not be used for gaining energy "except in vegetation." The first for whom the second law was of "combinatorial-statistical" nature was Ludwig Boltzmann; a view that we adopt here. (Boltzmann's life ended tragically with his suicide in Duino, Italy. This is just one facet adding up to a picture of darkness around that law; this is the expression of the "death drive" of physics, not its "Eros," such as Bell correlations and the fascination and promises they come with.)

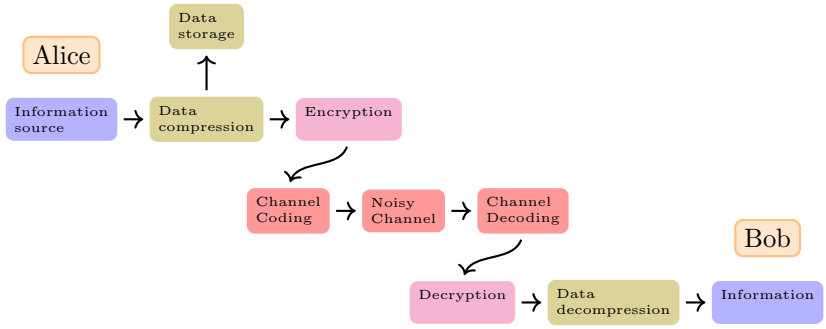
More serious than Kelvin's second thoughts related to photosynthesis is another life form, imagined by *Maxwell* in 1860, also in violation of the law: "Maxwell's demon." The latter is imagined to be a creature sitting at a separation wall in the gas that has a frictionless door, which the demon can open whenever a molecule moves towards that door from the left. In linear time in the number of particles of the gas, it will be "sorted", *i.e.*, compressed — revoltingly unfaithful to the statement of the second law. What is wrong with the argument? The understanding of this is one of the first success stories of the marriage between physics and *information theory*.

2.2 Information Theory

Information theory was developed in 1948 by *Claude E. Shannon* to determine the fundamental limits of signal-processing operations such as data compression on reliable storage and communicating. Now information theory is a field in its own right at the intersection of mathematics, statistics, computer science, physics, neurobiology, and electrical engineering. Since its inception it has been broadened to find applications in many other areas, including statistical inference, natural language processing, cryptography, neurobiology, the evolution and function of molecular codes, model selection in ecology, thermal physics, quantum computing, plagiarism detection, and other forms of data analysis.

2.2.1 Standard Model of Communication

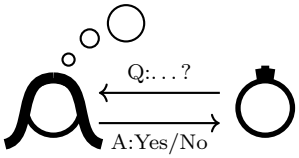
As information theory initially focused on communication, the scenario of two parties, usually referred to as Alice and Bob, sending one another some information, is common. A model usually contains *compression* to reduce the size of the data representing specific information, *encryption* to thwart attacks on information transfer, *channel coding* to introduce redundancy to protect against errors during transmission, as well as their reversing counterparts.



Claude Shannon considered parts of the scenario above separately in a formal manner. The key quantity do to so is a measure for information. How could we find a formal description of information? Shannon took an indirect approach considering "uncertainty" as opposed to "information." This leaves us with the problem of characterizing and measuring uncertainty.

2.2.2 The Game of 20 Questions

Imagine Alice knows a secret. Bob can retrieve information about the secret by asking (yes/no) questions.



How much information could he theoretically obtain by asking 20 such questions? The 20 (yes/no) answers Bob can get correspond to 20 pieces of binary information — 20 bits. These 20 bits allow Bob to distinguish 2^{20} different messages, as there are 2^{20} different sets of answers.

Conversely, one might ask, how many (yes/no) question are (at least) required to characterize an object in given set? Assume that Alice chooses an element $x \in \mathcal{X}$. Bob now wants to determine which element Alice has chosen. A bad strategy would be to ask for all elements $x \in \mathcal{X}$: have you chosen x? Until Alice eventually says yes. Instead, one could divide the set $\mathcal{X}$ into subsets of equal size and ask in which one the chosen element was, as shown in Figure 2.1. The number of questions to be asked is then again

$$\text{Number of questions} = \lceil \log_2 |\mathcal{X}| \rceil$$

These considerations motivate Hartley's formula for the uncertainty or entropy of a uniform random variable X over $\mathcal{X}$, with $P_X(x) = 1/|\mathcal{X}|$, is

$$H(X) = \log_2 |\mathcal{X}|$$

in units of bits. Suddenly, uncertainty is a function of a random variable. The connection to probability theory is explained below. For now, consider the special case of a uniform random variable over the set $\mathcal{X}$ just as another way of stating that Bob has no information about Alice's choice.

Why the logarithm? Apart from arguing with (yes/no) games, why should we use the logarithm for a measure of uncertainty? First, the uncertainty is supposed to be monotone with the size of the sample space $|\mathcal{X}|$. Second, the entity is supposed to be additive for joint distributions. For two sample spaces

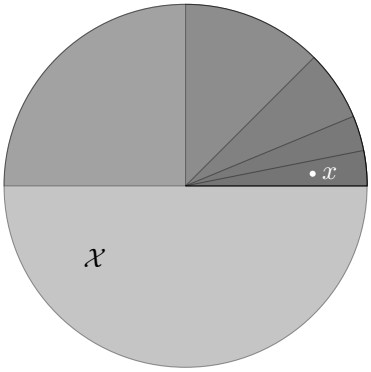


Figure 2.1: Strategy for determining an element that was previously chosen uniformly at random.

$\mathcal{X}$ and $\mathcal{Y}$ and a uniform random variable Z with $P(x \wedge y) = 1/(\mathcal{X} \cdot \mathcal{Y})$, Hartley's formula yields

$$H(Z) = \log_2(|\mathcal{X}| \cdot |\mathcal{Y}|) = \log_2 |\mathcal{X}| + \log_2 |\mathcal{Y}| = H(X) + H(Y)$$

2.2.3 Connection to Probability Theory

Bob's prior knowledge about Alice's choice can formally be described by a probability distribution. If Bob does not know anything about Alice's choice, each element in $\mathcal{X}$ was equally likely to be chosen by Alice. Thus Bob's knowledge corresponds to a uniform random variable over $\mathcal{X}$. The other extremal case is that Bob already knows which element Alice has chosen. This would then correspond to a deterministic probability distribution with $P_X(x) = 1$ for one particular $x \in \mathcal{X}$ and $P_X(x) = 0$ for all others. If Bob had known that Alice had chosen a word from an English dictionary, assuming that $\mathcal{X}$ contained all sequences up to 30 letters, then the probability distribution was uniform on the subset containing all dictionary words with less then 30 characters and zero for all other letter sequences. Bob might further be aware that Alice chooses her element according to the distribution of words in the English language. Frequent words like "the" or "and" are more likely to occur than for instance "supercalifragilisticexpialidocious."

Formal characterization of the entropy For a general (not necessarily uniform) random variable X, the measure of uncertainty, *i.e.*, the entropy, should correspond to the expectation value of the number of (yes/no) questions to find an element $x \in \mathcal{X}$ using an optimal strategy and combining asymptotically many realisations of X.

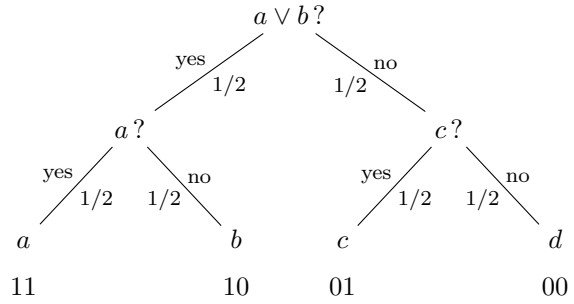


Figure 2.2: Strategy of determining a randomly chosen element in $\mathcal{X}$. Below the final result the binary representation of the sequence of answers is given. This could be regarded as the code representation of the particular element.

Example 2.2.1 (Entropy for uniform random variable). Let us consider a random variable X over $\mathcal{X} = \{a, b, c, d\}$ with $P_X(x) = 1/4$. A strategy similar to the one mentioned above yields the entropy $H(X) = 2$. It can be shown that the strategy we have employed here is optimal.

Example 2.2.2 (Entropy for non-uniform random variable). If we change the distribution of the random variable, the strategy used above is not optimal anymore. So, let us modify the probabilities to

$$P_X(a) = \frac{1}{2}\,, \quad P_X(b) = \frac{1}{4}\,, \quad P_X(c) = P_X(d) = \frac{1}{8}$$

Now, it is better to ask whether a is the right solution right away, as in half of the cases this will actually be right.

The expectation value of the number of question is slightly more complicated compared to the previous case

$$\begin{aligned} E[\#\text{ questions}] &= \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{8}\cdot 3 + \frac{1}{8}\cdot 3 \\ &= \left(\frac{1}{2}\right)^1 \cdot 1 + \left(\frac{1}{2}\right)^2 \cdot 2 + \left(\frac{1}{2}\right)^3 \cdot 3 + \left(\frac{1}{2}\right)^3 \cdot 3 \\ &= \sum_{x\in\mathcal{X}} P_X(x) \log_2\left(\frac{1}{P_X(x)}\right) \\ &= \frac{7}{4} < 2\,. \end{aligned}$$

Again, the strategy turns out to be optimal, though the proof is not given here. The entropy is smaller than in the uniform case. This fits the intuition that the uniform distribution corresponds to the least knowledge or the largest

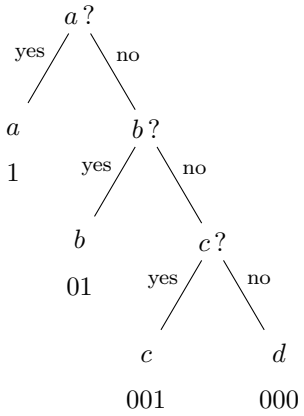


Figure 2.3: Questioning strategy for a non-uniform probability distribution. Again the binary codes are given below.

uncertainty. Deviating from the uniform distribution gives more and more information. The extreme case is finally the deterministic one, with one value occurring with probability 1, whereas all the others never happen at all. Then we do not even need one single question and the entropy is zero.

In both examples, the number of questions to determine a particular element corresponds to the length of the codeword representing the element in binary digits, *i.e.*, bits. As in the last example, the optimal length of the codeword is given by

$$l_c(x) = \log_2\left(\frac{1}{P_X(x)}\right) = -\log_2(P_X(x))\,.$$

Thus, we can now formally define the entropy.

Definition 1 (Entropy)**.** The entropy of a random variable X over $\mathcal{X}$ with a distribution $P_X(x) = p_x$, is given by

$$H(X) = E[-\log_2 P_X(x)] = -\sum_{x\in\mathcal{X}} p_x \log_2 p_x\,.$$

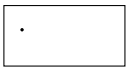
In case of a uniform distribution, the entropy of a random variable is the size of its range. This can be carried over to thermodynamics: "The entropy of a macrostate is the logarithm of the number of microstates corresponding to it." The *microstate* of a physical system, *e.g.*, a gas, specifies the position and momentum of each of the molecules. A *macrostate*, on the other hand, is simply a *set of microstates*, typically with a short description, such as:

"Helium gas at temperature 300 K in a cubic container of volume 100 liters and pressure 2 bar." So if all microstates in a macrostate "look roughly the same" — this latter condition is the translation of the uniformity condition for Hartley's formula above — we have[2]

$$H(\text{macrostate}) \approx \log_2(\#\text{ of corresponding microstates}) \ .$$

The "physical entropy," as opposed to the information-theoretic one, is usually called S and has an additional factor of $k \ln 2$, where k is Boltzmann's constant ($\approx 1.38 \cdot 10^{-23}$ Joule/Kelvin) and $\ln(2)$ is a common view at the border between nature and abstraction since 2 is a logical but not a natural constant.

Let us look at some examples. In the first one, we consider a "one-molecule gas" in a container of volume V, where the first macrostate corresponds to the particle being *anywhere in the container*, where for the other, smaller macrostate, the single particle is in the left half of the container.

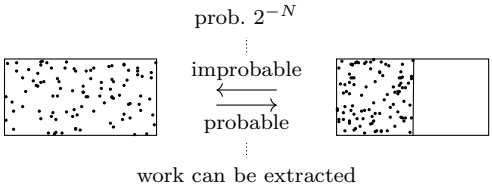


Two remarks: Obviously, not all microstates in a macrostate look the same; for instance, it is probably likelier that the particle is close to the middle of the container than to one of its corners. Second, we cannot simply *count* the number of possible positions since it is infinite for both macrostates. Indeed, there also exist definitions of entropy for the continuous case, but we do not need that here as long as we are solely interested in entropy *differences*: Let us assume that we distinguish different microstates only up to some fixed finite precision, *i.e.*, we put a grid of a certain "density" into the respective volumes, and the particle's position is described by such a grid point.

Then — whatever that density —there are exactly *twice as many* microstates corresponding to the first marcostate as to the second; the entropy difference is *one (bit)* $\Delta H = -1$, or, renormalized, $\Delta S = -k \ln 2$.

[2]There is something strange about this definition, namely, the apparent arbitrariness of how to assemble microstates together into macrostates. For instance — extremal example — if each macrostate corresponds to exactly one microstate, the entropy is always zero. Thus, the second law holds then, but it does not mean much. Maybe the macrostates correspond to *all we know* about a system. But then, the second law talks about *our knowledge* rather than the state of the system itself. Maybe we would like the macrostate to have a short description, as in the example above. This may work for equilibria — but is the second law not much more general?

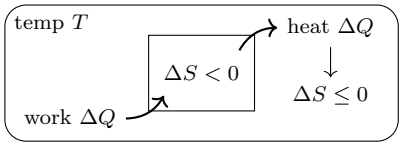
Let us look again at the gas with N molecules.



The entropy difference here is

$$\Delta S = -Nk \ln 2 \ .$$

We can say that the transition that is *probable* can also be said to allow for *gaining work*. Accordingly, we can of course enforce the "improbable" transition, for instance, by pushing a piston. The gas then ceases to be a closed system, so there is no contradiction. The price we have to pay for the compression is a certain amount of free energy. The contradiction to the second law is then avoided by that amount of energy being dissipated as heat to the environment, increasing the entropy around the gas container.



In the gas, the entropy decreases. A certain amount of work ΔQ is invested to force that decrease, which is then dissipated as heat to the environment. It is that heat dissipation that, after thermalization, is responsible for the compensation — in the form of an entropy increase in the environment — of the entropy decrease in the gas. Quantitatively, the entropy increase in a heat reservoir of temperature T, which receives an amount of ΔQ of heat energy at most equal to a value proportional to $1/T$, and Boltzmann's constant is defined such that we have

$$\Delta S \leq \frac{\Delta S}{T} \ .$$

Taken together, the minimal investment in terms of free energy to enforce that entropy decreases is, hence,

$$\Delta Q \geq \Delta S \cdot T \ .$$

Let us switch back to the example of the one-molecule gas. If we interpret the molecule as storing one bit, 0 if it is on the left and 1 on the right, then forcing the molecule to the left corresponds to the *erasure* of that bit, in the

sense that its state is 0 at the end, whatever it was before. The price for this erasure is then $kT \ln 2$: This fact — together with the claim that it is irrelevant by *what* physical system the bit is stored as long as it is stored by *some* physical system — is called *Landauer's principle.*

Landauer's principle. Erasing a bit requires

$$kT \ln 2 \quad (\approx 3 \cdot 10^{-21} J \text{ at room temperature})$$

free energy, which must in the process be dissipated as heat to the environment.

Example 2.2.3. Erasing a 80 GB iPod at room temperature.

$$\begin{aligned} \Delta Q &= 80 \cdot 2^{30} \cdot 2^3 kT \ln 2 \\ &\approx 2 \cdot 10^{-9} J \ . \end{aligned}$$

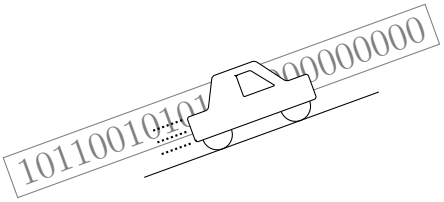
Example 2.2.4. Human body; brain: per discharge $10^{11} kT$. DNA replication: $20 - 100 kT$.

Example 2.2.5. Modern computer; transistor density: $10^7/\text{cm}^2$, frequency: $4 \cdot 10^{10}/s$, Binary erasure operation / tact: 0.1. Then the theoretical lower bound for power consumption and heat dissipation is $10^{-5} W/\text{cm}^2$. If this is extrapolated, assuming the validity of Moore's law, until 2030, we get $5 W/\text{cm}^2$ — way more than a stove.

The "thermodynamic wall" is today considered to be the most serious obstacle to Moore's law — even before quantum effects. If Landauer's principle is turned upside down, then a more positive statement results: Certain pieces of information, such as the string $000 \cdots 00$, have a work value.

2.3 The Converse of Landauer's Principle

The inverse process of erasure, let us call it *randomization*, allows for *gaining* free energy.



More precisely, the work value of the all-0-string of length N — however it is presented physically — is $kTN \ln 2$: This amount of environmental heat energy can be transformed into work.

Are there other strings with work value? The string consisting of the first N digits of the decimal expansion of π has work value (essentially) $kTN \ln 10$; the

reason is that it has a short description, *i.e.*, there exists a short program for a universal Turing machine to compute the string. This offers the possibility of a logically reversible computation between the given string and the all-0-string of the same length. We will see later that this means they have the same work value.

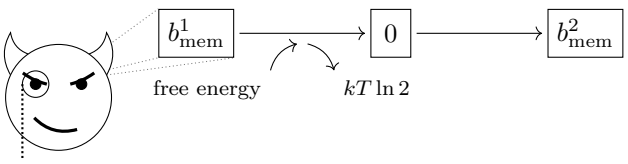
The given example already insinuates that the work value is directly connected to *reversible data compression*: Exactly those strings have work value that can be compressed to a shorter length in a lossless fashion. For instance, a string of length N resulting from flips of an unfair coin with probabilities p and $(1-p)$ has, most likely, work value $(1-h(p))kTN\ln 2$.

Accordingly, r copies of the same, perhaps incompressible, string of length N have work value at least $(r-1)kTN\ln 2$.

This example, for $r=2$, implies that if you have a car park of 2^N cars, each of which has one of the possible N-bit-strings hard-coded in it, then every string R can serve as fuel — for the right car, namely, the one that corresponds to just the string R. From the two copies of R, an amount of $kTN\ln 2$ can be generated, and another "random" string R' results, for which again there is a suitable car, and so on: We have constructed a *perpetuum mobile* of the second kind — or have we? Of course, we have not. And the explanation *why* we have not is exactly the same as for the problem of Maxwell's demon.

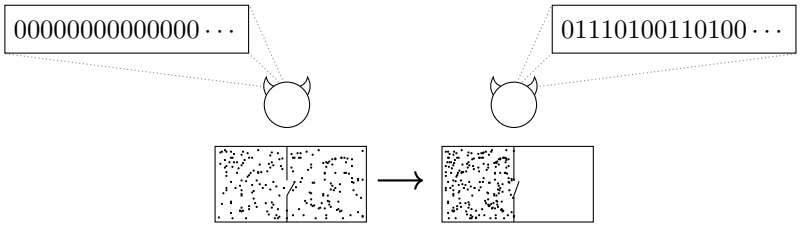
2.4 Bennett's Solution to the Problem of Maxwell's Demon

In the face of Landauer's principle, the paradox disappears: The demon must erase all the information that has accumulated in its brain during the sorting procedure, and the necessary heat dissipation exactly compensates for the entropy decrease, *i.e.*, the order created by the demon. More specifically, the demon must have an internal state depending on its observations and guiding its actions. In the case where the demon only has a one-bit-memory, this cell must be erased in every step; not that *overwriting* means erase and then write: Information gets forgotten.

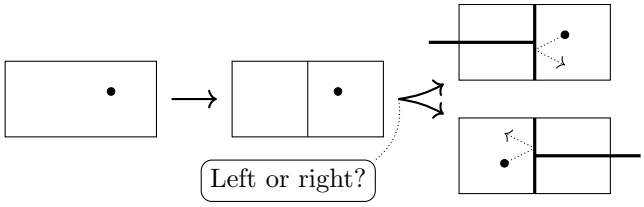


The argument gets even simpler when the demon has a large memory, which we assume to be in the all-0 state before the sorting procedure. Thereafter, the brain is filled with results of observations (in some sense: remembering

the original state of the gas that was forced to be forgotten by the gas itself through the demon's adaptive actions).



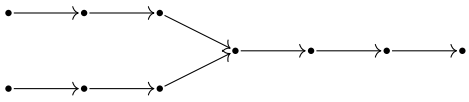
The fact that redundancy allows for gaining free energy can — for instance, where a demon has access to a second copy of some "random" string R out there — be read in the sense that "knowledge is energy." An early variation of this theme goes back to *Leo Szilard*, also the inventor of the nuclear chain reaction: If a one-molecule gas in a container is known to be on, say, the left half of the container, then this knowledge can be transformed into free energy (again, more precisely: used to transform environmental heat energy into free mechanical energy).



Symmetric constructions can also be imagined here. They do, however, not constitute a *perpetuum mobile* of the second kind since the partition must, after the work extraction, be put back to the middle — by a "demon" who must then erase its brain at the corresponding thermodynamic cost.

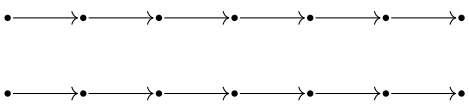
2.5 Reversible Computing

Motivated by the considerations in the previous section, we turn back to computing and ask, in particular, the question whether it can be made *logically reversible*.



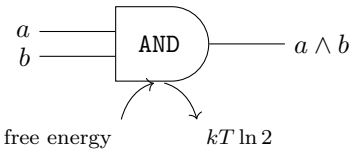
If, in the course of the computation, information is lost about "which branch the computation came from," then free energy $kT \ln 2$ must be invested, which is dissipated as heat to the environment: In other words, the

logical irreversibility of a computation ("information is lost") implies its *thermodynamic* irreversibility ("free energy is lost"). From that point of view, it appears advantageous if a computation does not have such collisions.

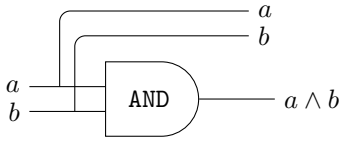


We will see that it is in fact possible to carry out any logically reversible computation in a thermodynamically reversible way in principle — Landauer's principle is in this sense "tight."

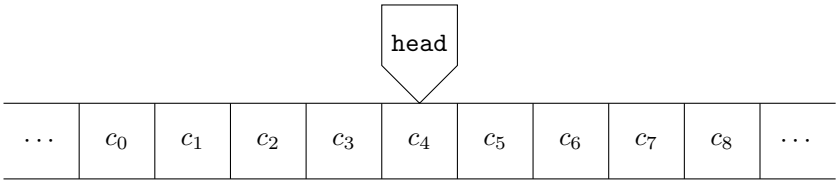
Let us consider a concrete example of a gate, the logical AND.



Obviously, it is logically irreversible since the output does not allow for completely reconstructing the input. Can we modify the gate in order to render it logically reversible? A first idea is to keep the inputs and have them be part of the output.

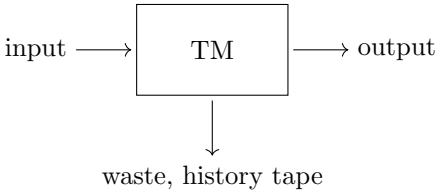


More generally, a general computation by a logically irreversible Turing machine — which can still be deterministic, as we saw at the beginning of this chapter — can be made logically reversible if the machine is additionally given a "history tape" for storing the entire path of the computation. This was the starting point of Bennett's idea to make any computation reversible.

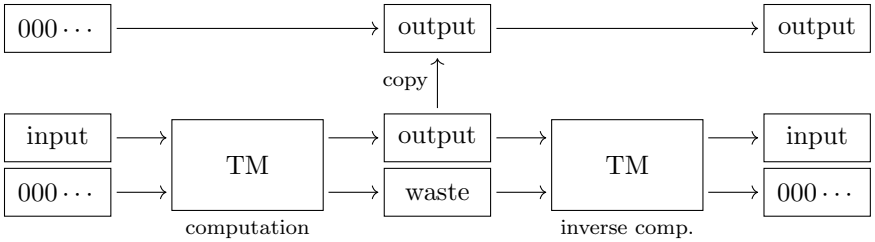


If the machine writes the entire "history" of the computation to that extra tape, then the problem is not solved, or let us say: The solution is not "sustainable." The reason is that the original state of the history tape, say

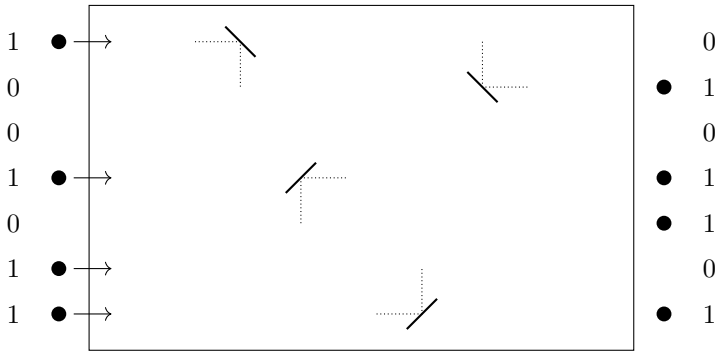
the all-0-string, is lost and replaced by the waste that has piled up — just as in Maxwell's demon's brain. (If the history tape was not actually filled with all-0s but with, say, a "random" string, then that has to be overwritten, which already wastes free energy as heat into the environment, and it becomes irreversible.)



Bennett's idea was to rid of that waste in on "orderly fashion," *i.e.*, to *uncompute* instead of erase it: The computation is done, including the history tape, the output is copied on some extra bit positions, and then the computation is undone step by step in reverse order.



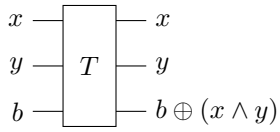
We have hence seen that any computation can be made logically reversible. Can every logically reversible computation be carried out in a thermodynamically reversible way? An affirmative answer to this was given by *Fredkin* and *Toffoli* in the form of the "Ballistic computer": Elastic collisions of balls on a billiard table can carry out every computation by a Turing machine — as long as no information is lost. (The latter is impossible due to the time-reversal symmetry of the laws of classical mechanics.)



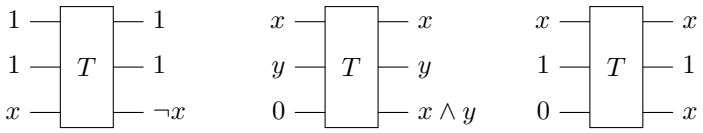
Given this set of encouraging results, people have been motivated to systematically study reversible computing, and in particular what could be its basic building block (such as the NOT and AND gates for irreversible computation).

2.6 The Toffoli Gate

We have had a look at an AND gate before, and observed that it can be made logically reversible if the inputs are not "thrown away" but stored and made part of the output. There is then an extra wire, the "output wire," yielding the actual result of the gate. Since we also want to keep track of the original state of the particular physical degree of freedom representing that output, we already have the output wire as part of the input; in fact, reversible gates (and thus: circuits) always have the same number of input and output wires. Finally, changing the state of the input state of the output wire should also change the state of that wire at the output; otherwise, information would be lost again. Altogether, we get the following made-reversible AND: the *Toffoli gate*.



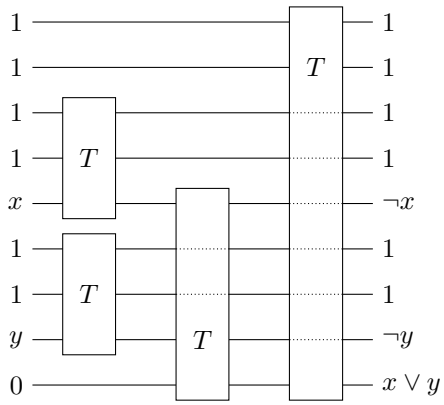
The Toffoli gate is universal: Any circuit can be translated into one using only such gates (and, in particular, no fan-outs). This follows from the possibility to get, from one Toffoli each, the NOT, the AND, as well as the FAN-OUT.



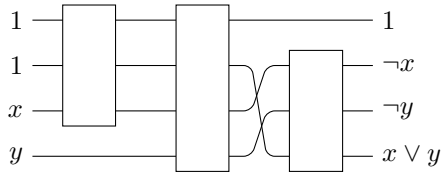
In general, additional (constant) inputs are used, and additional outputs produced. Let us, for example, compute an OR. If we apply de Morgan's formula, then we get

$$x \vee y = \neg(\neg x \wedge \neg y) \ ,$$

and it is obviously possible with *four* Toffolis.

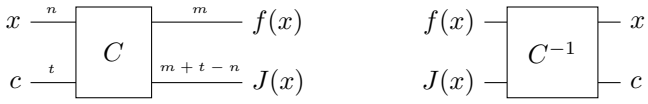


This can be improved in two ways: First of all, we can recycle the constants (1 in this case), and we do not need the fourth Toffoli if the last input to the third Toffoli is changed to 1. Then the final negation is for free.

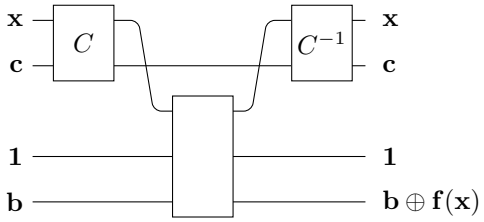


Any arbitrarily large circuit, consisting of irreversible gates including AND and OR as well as FAN-OUTs can now be translated into a reversible circuit using only Toffolis. The circuit then has the same number of input and output wires. In general, the reversible circuit, however, besides the actual function input, also needs a number of "constants" to be put in. On the respective wires, at the end of the computation, there will then be intermediate results of the computation — resembling in a sense the content of the reversible Turing machine's history tape, also reflecting the entire path leading up to the desired output result. And just as in that latter example, the reversible-making is not finished or not "sustainable," as long as the generated junk (*i.e.*, these intermediate results replacing the constants) is not gotten rid of in an orderly way. And the analogy goes even further: It *can* be gotten rid of — and how this works is exactly how Bennett did it: It can be *uncomputed* step by step after the computation and the copying of the output of interest. (Note that, in a *quantum computer*, this *getting rid of extra degrees of freedom* is even more important, since losing even only one single Qbit of a computation destroys all decoherence: The computation breaks down.

Let us now assume we have a *reversible circuit with junk* computing a function f from m to n bits. Note first that the circuit can be used in either direction.



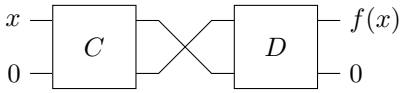
Bennett's trick for uncomputing the junk then looks as follows: First, C is applied, then the output of interest is copied, using one Toffoli per bit to realize a FAN-OUT, and then C^{-1} is used for uncomputing the junk. This way, we get the required constants back, and they do not even have to be taken into account in the input/output behavior of the circuit.



We can conclude that the best reversible circuit is, hence, only roughly *half as efficient* as the best *irreversible* circuit: The loss in efficiency is only a constant factor of 2. Note that, in the resulting circuit, the input x appears again as part of the output; after all, this is necessary in general to guarantee the logical reversibility of the whole gate. There is, however, a situation for which this is not true: If f is a bijective function. Can there then also exist a reversible circuit that takes only x as input and that outputs *only* $f(x)$? If yes, what about the efficiency? Does reversibility again only cost as a factor of 2?

According to the above, there exist classical circuits for the function f as well as for its inverse f^{-1}, let us call then C and D. (Note that D actually allows for computing the inverse of f — unlike C^{-1}, which requires the answer, x, already as an input.)

Now, these two circuits can be connected in the following way to get the desired simple circuit for f *not* again outputting the input.



Note, however, that this circuit *may sometimes be necessarily much less efficient than the best irreversible circuit for* f. In fact, the latter is the case if and only if f is a "one-way function," a central notion in cryptography: A *one-way function* is a bijective function that can be computed efficiently, but for which no efficient inversion algorithm exists. Clearly, a function for which the irreversible circuit is efficient *cannot* be one-way since that circuit can be used on both ways. (Again, not that this does not apply for those reversible circuits for which x is again part of the output.) On the other hand, a function for which the reversible circuit is necessarily inefficient cannot have an efficient inversion algorithm because of the construction shown.

All in all, we have here a fascinating example where *thermodynamics meets cryptography.*

Chapter 3

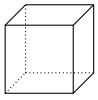
Key Experiments and Postulates of Quantum Physics

The experiments previously mentioned to illustrate the particular effects of quantum mechanics historically did not mark the beginning of the theory. Instead, it was the *UV catastrophe* of black-body radiation and the photoelectric effect that lead to the development of quantum physics. We will briefly discuss the two corresponding experiments and how they broke with classical expectations before working our way towards the formal basis of quantum mechanics.

3.1 Black-Body Radiation

In classical statistical mechanics, the law of equipartition expresses the idea that, in thermal equilibrium, energy is usually shared among all possible corresponding microstates. Each degree of freedom carries the same average energy $E = kT/2$, where k is Boltzmann's constant. Assuming that heat is transferred by electromagnetic waves, the equipartition law leads to problematic consequences.

Let us imagine an idealized cubic vacuum with side length l that absorbs and emits all radiation frequencies. The cube is in thermodynamic equilibrium with its environment, a heat bath of temperature T.



The thermal electromagnetic radiation within this body can be thought of as standing waves within the body. In each of the three spatial dimensions, the standing wave is characterized by its number of node points. Generally, a standing wave is then a superposition of three standing waves, each corresponding to one spatial direction. Thus, such a wave can be described by a three-dimensional vector of positive integers $\vec{n} = (n_1, n_2, n_3) \in (\mathbb{N}_{>0})^3$.

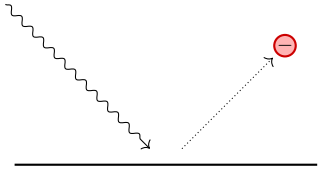
The possible waves correspond to integer vectors in the first octant of the $\vec{n}$-space. The number of waves in a shell of radius $|\vec{n}|$ scales as the surface of the shell and thus as $|\vec{n}|^2$. Therefore, classically the energy density should have scaled as well with $|\vec{n}|^2$. Likewise, the energy density scales with the square of the frequency $\omega = c\pi |\vec{n}| / l$. This growth of the energy density is usually referred to as the *ultraviolet catastrophe*. It is absurd in many respects, and in particular, everyone in a classroom at room temperature would immediately drop dead due to the intensity of X-rays.

While for lower frequencies the quadratic dependence of the spectral radiance on the frequency—the *Rayleigh-Jeans law*—matches experimental findings, this is not the case for higher frequencies: The spectral radiance exponentially *decreases* again for higher frequencies.

Quantization Max Planck assumed the radiation energy to be absorbed and emitted merely in integer multiples of $\hbar\omega$, where $\hbar = 1.054 \times 10^{-34} Nms$. The probability of emission or absorption decreases exponentially in $\hbar\omega/kT$. In particular, the probability falls of exponentially for large frequencies—in accordance with experimental findings.

3.2 Photoelectric Effect

In 1887, *Heinrich Hertz* examined the emission of electrons from a metal surface if light is shone onto it. Classically, the expectation was that the velocity of the emitted electrons is faster depending on the *intensity of the light* but independent of the color, *i.e.*, the frequency of the light.



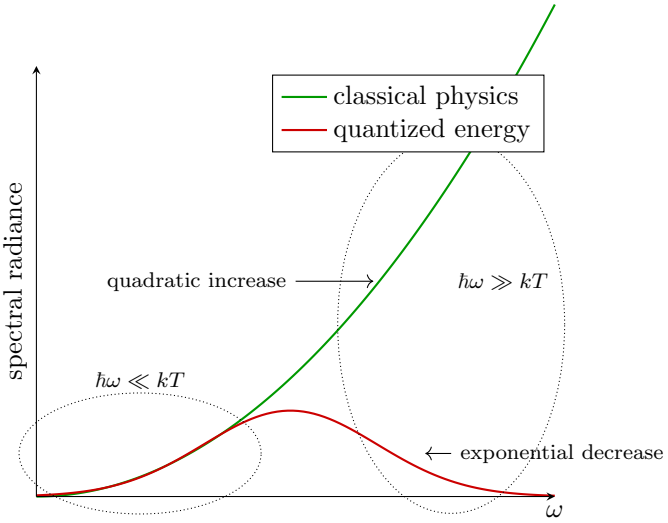
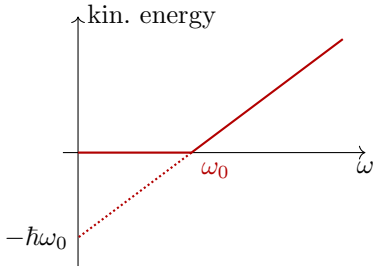


Figure 3.1: According to the Rayleigh-Jeans law, the spectral radiance grows quadratically in the frequency, thus leading to the ultra-violet catastrophe. In Planck's law, in turn, the spectral radiance decreases exponentially for large frequencies.

Hertz, however, found that the intensity merely changed the *number* but not the velocity of the emitted electrons. Instead, increasing the *frequency* of the light increased the velocity of the electrons. Below a certain threshold frequency ω_0, there were no emitted electrons, whatever the intensity or the duration of shining the light onto the surface.



How could it be that light—*i.e.*, electromagnetic waves—does not provide enough energy for the emission of electrons even if the intensity is increased? The energy of a wave depends on its amplitude and *not its wavelength*. The energy should then accumulate until the electron would finally be emitted.

In 1905, Albert Einstein provided an explanation for the observed behavior by regarding light as a quantized packet of a certain, frequency-dependent energy—by attributing particle properties to light. Light comes then in packets of $\hbar\omega$ energy. If the frequency is lower than ω_0, then the energy lies below the energy $W = \hbar\omega_0$ needed to remove the electron from the metal. The kinetic energy of the electron is then $E_{kin} = \hbar\omega - W$.

If light is to be regarded as particles, so-called *photons*, then how do we make sense of the wave properties of light? If, for example, light with a certain polarization is shone onto a polarizing filter rotated by 45° with respect to the initial polarization, then the intensity of the light reduced by a factor 1/2. This cannot be explained anymore by a reduced amplitude. Also, the color of the light does not change. The reduction of intensity can, however, be understood as a probability measure: With the probability 1/2, the photon is absorbed by the filter. This shows how direct and small the step is from quantization to a probabilistic interpretation — the two most marking features of quantum theory.

The group of experiments discussed shows the other side of a fundamental dualism in quantum mechanics: Not only do particles behave like waves, also waves behave like particles.

3.3 Wave-Particle Dualism

We now examine the wave-particle dualism using the example of a plane wave.

$$\psi(\vec{x}, t) = C \cdot e^{i(\vec{k}\cdot\vec{x} - \omega t)} \ ,$$

where $\vec{k}$ is the wave vector, perpendicular to the wave fronts which relates to the wavelength as $\lambda = 2\pi/|\vec{k}|$. The period of the wave, *i.e.*, the time that elapses between the passage of two wave fronts, is $\Delta t = 2\pi/\omega$. The phase velocity $v = \omega/\left|\vec{k}\right|$ is the velocity of the wave front.

> **Interlude: Hilbert spaces**
>
> The state spaces of quantum-mechanical systems are generally so-called *Hilbert spaces*, *i.e.*, complete complex inner-product spaces. Quantum *informatics* happens in *finite*-dimensional Hilbert spaces, whereas they can be uncountably infinite-dimensional in general. An important result states that, for a given dimension, there exists *exactly one* Hilbert space.

The wave function $\psi(\vec{x}, t)$ is an element of an infinite-dimensional Hilbert space. The inner product is defined as

$$(\psi_1, \psi_2) = \int_{-\infty}^{+\infty} \psi_1^*(\vec{x}, t)\psi_2(\vec{x}, t)d\vec{x}\,.$$

For the integral to be well defined, we have to restrict to the square-integrable functions. In order to allow for a normalization ψ, the factor C would have to be an envelope that makes the integral finite. This yields a wave packet that is then again a superposition of plane waves.

We can now replace $\vec{k}$ and ω in the plane wave equation: On the one hand, *Louis de Broglie* associated the momentum of a particle with its quantum-mechanical wave vector as $\vec{p} = \hbar\vec{k}$. On the other hand, the momentum can be related to the energy $E = \hbar\omega$ as in classical mechanics as

$$E = \hbar\omega = \frac{\vec{p}^2}{2m}$$

and, thus, $\omega = \vec{p}^2/2m\hbar$. The plane wave can then be written as

$$\psi(\vec{x}, t) = C \cdot e^{i(\vec{p}\cdot\vec{x} - frac p^2 2mt)/\hbar}\,.$$

So, what is now the *time evolution* of the plane wave $\psi(\vec{x}, t)$, *i.e.*, the function

$$\psi(\vec{x}, 0) \mapsto \psi(\vec{x}, t)\,?$$

Let us examine the partial derivative of the plane wave:

$$\begin{aligned}
\frac{\partial}{\partial t}\psi(\vec{x}, t) &= \psi(\vec{x}, t) \cdot \left(\frac{i}{\hbar} \cdot \left(-\frac{p^2}{2m}\right)\right) \\
&= (p_1^2 + p_2^2 + p_3^2)\psi(\vec{x}, t) \cdot \left(-\frac{i}{2m \cdot \hbar}\right) \\
&= -\hbar^2 \left(\left(\frac{\partial}{\partial x_1}\right)^2 + \left(\frac{\partial}{\partial x_2}\right)^2 + \left(\frac{\partial}{\partial x_3}\right)^2\right)\psi(\vec{x}, t) \cdot \left(-\frac{i}{2m \cdot \hbar}\right) \\
&= -\hbar^2 \Delta\psi(\vec{x}, t) \cdot \left(-\frac{i}{2m \cdot \hbar}\right)\,.
\end{aligned}$$

This directly corresponds to the Schrödinger equation of a free particle

$$i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\Delta\psi\,.$$

Interlude: Hermitian and unitary operators

The adjoint of a linear operator is defined through the inner product as follows:

$$(\psi_1, A\psi_2) = (A^*\psi_1, \psi_2)\,.$$

A linear operator is self-adjoint or Hermitian if and only if $A^\dagger = A$.

A linear operator is called *unitary* if and only if $U \cdot U^\dagger = \mathbb{1}$.

There are two important properties of the Schrödinger equation: It is linear, and it preserves the inner product. The first property means: Any linear combination of solutions of the Schrödinger equation

$$\psi(\vec{x}, t) = \alpha\psi_1(\vec{x}, t) + \beta\psi_2(\vec{x}, t) \qquad \forall \alpha, \beta \in \mathbb{C}$$

again yields a solution. The second property means: A solution can be written as $\psi(\vec{x}, t) = U(t)\psi(\vec{x}, 0)$. Then, the norm square can be written as

$$|\psi(t)|^2 = (U(t)\psi_0, U(t)\psi_)) = (U^\dagger(t) \cdot U(t)\psi_0, \psi_0) = (\psi_0, \psi_0)\ ,$$

where the latter equality holds for all ψ_0 if and only if $U^\dagger(t) \cdot U(t) = \mathbb{1}$. Thus, the time evolution operator is unitary. This also implies: The time evolution in quantum mechanics is reversible.

3.4 Observables

What is the expected position of a particle in a given quantum state ψ? We can consider $|\psi(\vec{x}, t)|^2 = \overline{\psi}(\vec{x}, t) \cdot \psi(\vec{x}, t)$ as the probability density for finding the particle in position $\vec{x}$ at time t. Then the expectation value becomes

$$\begin{aligned}
E_\psi[\vec{x}] &= \langle\, \vec{x}\, \rangle_\psi \\
&= \int d^3x\ \ |\psi(\vec{x}, t)|^2\, \vec{x} \\
&= \int d^3x\ \overline{\psi}(\vec{x}, t) \cdot \psi(\vec{x}, t)\vec{x} \\
&= \int d^3x\ \overline{\psi}(\vec{x}, t) \cdot \vec{x} \cdot \psi(\vec{x}, t) \\
&= \int d^3x\ \overline{\psi}(\vec{x}, t) \cdot A(\psi)(\vec{x}, t) \\
&= (\psi, A\psi)\ ,
\end{aligned}$$

where $A : \psi \mapsto \vec{x} \cdot \psi$ is the *position operator*. In quantum mechanics, any measurable quantity corresponds to a self-adjoint operator, a so-called *observable*.

We can now repeat the consideration above for the expectation value of the momentum of the particle. Bear in mind the plane wave

$$\psi(\vec{x},t) = C \cdot e^{i(\vec{p}\cdot\vec{x} - frac p^2 2mt)/\hbar} .$$

We can then expand the expected momentum as

$$\begin{aligned}
\langle\, \vec{x}\,\rangle_\psi &= \int d^3x\ \overline{\psi}(\vec{x},t)\cdot\psi(\vec{x},t)\cdot\vec{p} \\
&= \int d^3x\ \overline{\psi}(\vec{x},t)\cdot\vec{p}\cdot\psi(\vec{x},t) \\
&= \int d^3x\ \overline{\psi}(\vec{x},t)\cdot\frac{\hbar}{i}\vec{\nabla}\psi(\vec{x},t) \\
&= (\psi, frac\hbar i\vec{\nabla}\psi) \ ,
\end{aligned}$$

where $A = \hbar\vec{\nabla}/i$ is the momentum operator.

Similarly for the expected energy we obtain

$$\begin{aligned}
\langle E\rangle_\psi &= \int d^3x\ \overline{\psi}(\vec{x},t)\cdot\psi(\vec{x},t)\cdot E \\
&= \int d^3x\ \overline{\psi}(\vec{x},t)\cdot i\hbar\frac{\partial}{\partial t}\psi(\vec{x},t) \\
&= (\psi, i\hbar\frac{\partial}{\partial t}\psi) \ .
\end{aligned}$$

There appears the *correspondence principle*: Any measurable quantity in classical mechanics can be translated to quantum mechanics by using the corresponding Hermitian operator. If we consider the energy in classical mechanics, $E = \vec{p}^2/2m$, then we can replace the left side with $E = i\hbar\partial/\partial t$ and the right side with $\left(\hbar/i\vec{\nabla}\right)^2/2m$, and we return to the Schrödinger equation for a free particle

$$i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}(\vec{\nabla}\cdot\vec{\nabla})\psi = -\frac{\hbar^2}{2m}\Delta\psi \, .$$

The general, time-independent Schrödinger equation is

$$\frac{\partial}{\partial t}\psi = \frac{i}{\hbar}H\psi \, .$$

In the special case of a free particle, the *Hamiltonian* is

$$H = \frac{\hbar^2}{2m}\Delta \ .$$

The Hamiltonian is again a Hermitian operator. The general solution is then

$$\psi(\vec{x},t) = e^{iHt/\hbar}\psi(\vec{x},0) \ ,$$

where the exponential is defined by its series expansion

$$e^A = \sum_n \frac{A^n}{n!} \ .$$

This yields the time evolution operator, $U(t) = e^{iHt/\hbar}$. We then have

$$\begin{aligned} U(s) \cdot U(t) &= U(s+t) \ , \\ U(0) &= \mathbb{1} \ , \\ U(t)^\dagger &= U(t)^{-1} = U(-t) \ , \end{aligned}$$

as can be seen from

$$e^{iHt/\hbar} \cdot \left(e^{iHt/\hbar}\right)^\dagger = e^{iHt/\hbar} \cdot e^{-iH^\dagger t/\hbar} = e^{i(H-H^\dagger)/\hbar} = \mathbb{1} \, .$$

Thus, the exponential of a Hermitian operator yields a unitary operator.

3.5 Postulates of Quantum Theory

The postulates of quantum mechanics form the axiomatic basis of the theory. They summarize and formalize the discussion in the previous sections. After introducing the postulates of the theory for pure states, corresponding to normalized vectors in a Hilbert space, we extend this to statistical mixtures of such states, corresponding to *density matrices*.

3.5.1 The State

In quantum mechanics, a system — for instance, an electron, a photon, or an atom — is assigned a *normalized state vector* in a complex Hilbert space,

$$\psi(\vec{x}, t) \in \mathcal{H} \quad \|\psi\| = 1 \, .$$

If the inner product of the Hilbert space is defined as

$$(\psi, \varphi) := \int d^3x \ \overline{\psi(\vec{x}, t)} \cdot \varphi(\vec{x}, t) \ ,$$

then the normalization condition reads as

$$\|\psi\|^2 = (\psi, \psi) = \int d^3x \ \overline{\psi(\vec{x}, t)} \cdot \psi(\vec{x}, t) = 1 \ .$$

Superposition Any convex combination of state vectors

$$\alpha\psi_1 + \beta\psi_2 \in \mathcal{H}, \quad \alpha, \beta \in \mathbb{C} \text{ with } |\alpha|^2 + |\beta|^2 = 1$$

is again a vector associated with a state of the system—just as the superposition of two waves is again another wave. Correspondingly, this linear structure of the state space gives rise to interference effects and, generally, the wave characteristics of quantum mechanics.

3.5.2 The Time Evolution

The time evolution of a quantum state is governed by the Schrödinger equation

$$\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = H\psi \ ,$$

with $\hbar = 1.055{\cdot}10^{-34} Nms$ being the *reduced Planck's constant*. The Schrödinger equation is linear: Any linear combination of solutions to the differential equation constitutes also a solution.

We assume the Hermitian Hamilton operator, or Hamiltonian H, to be time-independent. Then, the solutions to the differential equation can be written as

$$\psi(t) = e^{iHt/\hbar}\psi(0) \ ,$$

with $\psi(0)$ being the state at some initial time t_0. Importantly, with the Hamiltonian H being Hermitian, the operator $e^{iHt/\hbar}$ is unitary and, thus, preserves the inner product. Consequently, the states $\psi(t)$ are normalized if the initial state $\psi(0)$ is.

$$\begin{aligned} &\text{If } (\psi(0), \psi(0)) = 1 : \\ &\quad (\psi(t), \psi(t)) = 1 \quad \forall t \end{aligned}$$

3.5.3 Observables

In quantum mechanics, measurable entities correspond to observables, that is, Hermitian operators A with $A^\dagger = A$. However, more generally, *any* Hermitian operator is an observable. Thus, the concepts of what are potentially measurable entities is both more abstract and more general. Previously, the expectation of an observable A, $\langle A\rangle_\psi = (\psi, A\psi)$ was mentioned. We now link this with actual measurement results.

The *spectral theorem* for finite dimensional linear operators states that an operator A has a spectral decomposition with real eigenvalues *if and only if*

the operator is Hermitian. In more formal terms, this reads as:

$$\begin{aligned} A = A^\dagger \;\Leftrightarrow \qquad & \text{There exist } \{\varphi_i\}_i \text{ with } (\varphi_i, \varphi_j) = \delta_{ij} \\ & \text{with corresponding projectors } P_{\varphi_i} \\ & \text{and } \lambda_i \in \mathbb{R} \\ & \text{such that } A = \sum_i \lambda_i P_{\varphi_i} \ . \end{aligned}$$

So an observable has a corresponding spectral decomposition. If one performs a measurement on a quantum system, one obtains one of the eigenvalues of the corresponding observable A as its result. The probability of measuring the value λ_i, one of the real eigenvalues of A, corresponding to an eigenvector φ_i when measuring a system in a state $\psi \in \mathcal{H}$ is

$$\begin{aligned} P(\lambda_i) &= \|P_{\varphi_i}\psi\|^2 = (\, P_{\varphi_i}\psi,\, P_{\varphi_i}\psi\,) \\ &= (\, (\varphi_i,\psi)\varphi_i,\, (\varphi_i,\psi)\varphi_i\,) \\ &= \overline{(\varphi_i,\psi)}\,(\varphi_i,\psi)\,\underbrace{(\varphi_i,\varphi_i)}_{=1} = |(\varphi_i,\psi)|^2 \ , \end{aligned}$$

where we have expanded the operator as $P_{\varphi_i}\psi = (\varphi_i,\psi)\varphi_i$ and used the sesquilinearity of the inner product.

Let us return to the expectation value mentioned above:

$$\begin{aligned} \langle A\rangle_\psi = (\psi, A\psi) &= (\psi, \sum_i \lambda_i P_{\varphi_i}\psi) \\ &= \sum_i \lambda_i (\psi,\, (\varphi_i,\psi)\varphi_i) \\ &= \sum_i \lambda_i \underbrace{(\varphi_i,\psi)\,(\psi,\varphi_i)}_{|(\varphi_i,\psi)|=P(\lambda_i)} \ . \end{aligned}$$

Consistent with the concept of the expected value from probability theory, we obtain the weighted sum over all possible results of the measurement corresponding to A.

3.5.4 Joint Systems and Composition

In classical mechanics, joint systems are described by combining the vectors of phase-space coordinates of the two by a Cartesian product. The Cartesian product, however, does not preserve the linear structure and thus the superposition principle of the subsystems. If $v \otimes w$ denotes the combination of two vectors $v \in \mathcal{H}_1$ and $w \in \mathcal{H}_2$, then this combination should have the properties

$$\begin{aligned} (v_1 + v_2) \otimes w &= v_1 \otimes w + v_2 \otimes w \ , \\ v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2 \ , \\ a(v \otimes w) &= av \otimes w = v \otimes aw \ , \end{aligned}$$

where $v, v_1, v_2 \in \mathcal{H}_1$, $w, w_1, w_2 \in \mathcal{H}_2$, and $a \in \mathbb{C}$. Then, any linear combination in any of the subspaces corresponds to a linear combination in the joint space. This is the essential characteristic of the *tensor product*.

Given two systems with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the joint system has a state space isomorphic to the tensor product of the Hilbert spaces $\mathcal{H}_A \otimes \mathcal{H}_B$. If the systems are in *pure* states $\psi_A \in \mathcal{H}_A$ and $\psi_B \in \mathcal{H}_B$, then their joint state is

$$\psi_A \otimes \psi_B \in \mathcal{H}_A \otimes \mathcal{H}_B \ .$$

However, any superposition of such product states is also a state in the joint Hilbert space, and these superposition states may not have a representation as a product. As is discussed subsequently, such states are called *entangled*.

3.5.5 Abstraction and Simplification

From now on we assume the Hilbert space to be finite-dimensional. Then, the Hilbert space is isomorphic to an n-dimensional complex vector space, $\mathcal{H} \cong \mathbb{C}^n$. With this assumption, the states can be expressed by coordinates with respect to some fixed basis. For $\varphi \in \mathcal{H} \cong \mathbb{C}^n$, we call the corresponding column vector "ket" of φ and represent it as

$$|\varphi\rangle = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{pmatrix} \ .$$

Similarly, the "bra" is the complex conjugate transpose of a ket, *i.e.*,

$$\langle\varphi| = |\varphi\rangle^\dagger = \begin{pmatrix} \overline{\varphi_1} & \overline{\varphi_2} & \cdots & \overline{\varphi_n} \end{pmatrix} \ .$$

With the convention above, the inner product can be written as

$$(\varphi, \psi) = \begin{pmatrix} \overline{\varphi_1} & \overline{\varphi_2} & \cdots & \overline{\varphi_n} \end{pmatrix} \cdot \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} = \langle\varphi|\psi\rangle \in \mathbb{C} \ ,$$

where $\cdot$ indicates matrix multiplication. This notation is commonly referred to as Dirac's Bra-ket notation.

The notation also proves useful in expressing the Hermitian observables. According the spectral theorem, the effect of an observable A on a vector ψ

may be expanded as

$$\begin{aligned} |A\psi\rangle &= \sum_i \lambda_i |P_{\varphi_i}\psi\rangle \\ &= \sum_i \lambda_i (\varphi_i, \psi)|\varphi_i\rangle \\ &= \sum_i \lambda_i \langle\varphi_i|\psi\rangle\, |\varphi_i\rangle \\ &= \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|\, |\psi\rangle \ . \end{aligned}$$

In the last step, we rearranged the bras and kets and obtained the common way to represent a projector as

$$P_{\varphi_i} = |\varphi_i\rangle\langle\varphi_i|$$

and thus the observable itself as $A = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$.

More generally, a linear operator on a finite-dimensional vector space can be represented with a fixed basis as a matrix. In Dirac's notation by abbreviating the basis vectors to their indices ($|\varphi_i\rangle =: |i\rangle$), this reads as

$$A = \sum_{k,l} \langle k|A|l\rangle\, |k\rangle\langle l| \ ,$$

where $\langle k|A|l\rangle$ is the matrix entry of the k-th row and the l-th column. The way in which it is written here is rather an expansion in a basis of the vector space of linear operators $\mathrm{End}(\mathcal{H})$, and $\langle k|A|l\rangle$ is the coefficient before the basis vector $|k\rangle\langle l|$, *i.e.*, a matrix with a one in the k-th row and the l-th column.

Again, we can reorder bras and kets to obtain

$$A = \sum_{k,l} |k\rangle\langle k|\, A\, |l\rangle\langle l| = \underbrace{\sum_k |k\rangle\langle k|}_{=\mathbb{1}}\ A\ \underbrace{\sum_l |l\rangle\langle l|}_{=\mathbb{1}} \ .$$

A common trick when calculating within the braket notation is to insert $\mathbb{1} = \sum_i |i\rangle\langle i|$ and then to rearrange the elements.

Exercise 1. As an exercise, you can confirm that a matrix $A \in \mathrm{End}(\mathcal{H})$ is invariant under change of basis. To do so, you may first expand A using a basis $\{|u\rangle\}_u$ of $\mathcal{H}$, and then insert twice the identity expressed in an arbitrary, different basis $\{|v'\rangle\}_v$.

The trace is an important linear map $\mathrm{End}(\mathcal{H}) \to \mathbb{C}$ defined as the sum over the diagonal elements of the matrix corresponding a linear operator $A \in \mathrm{End}(\mathcal{H})$

$$\mathrm{Tr}(A) := \sum_k \langle k|A|k\rangle \ .$$

Because this definition depends on the choice of a basis, we have to ensure that this choice is irrelevant for the trace. It can be shown that $\mathrm{Tr}(A \cdot B) = \mathrm{Tr}(B \cdot A)$. Therefore, for some unitary U, corresponding to a change of basis, we obtain

$$\begin{aligned}\mathrm{Tr}(U \cdot A \cdot U^\dagger) &= \mathrm{Tr}(U^\dagger \cdot U \cdot A) \\ &= \mathrm{Tr}(\mathbb{1} \cdot A) = \mathrm{Tr}(A) \ .\end{aligned}$$

More directly, using the Dirac notation, one can as well compute

$$\begin{aligned}\mathrm{Tr}(A) &= \sum_k \langle k|A|k\rangle \\ &= \sum_k \langle k| \sum_u |u'\rangle\langle u'| \, A \sum_v |v'\rangle\langle v'| \, |k\rangle \\ &= \sum_{k,u,v} \langle k|u'\rangle \, \langle u'|A|v'\rangle\langle v'|k\rangle \\ &= \sum_{u,v} \underbrace{\sum_k \langle v'|k\rangle\langle k|u'\rangle}_{=\langle v'|u'\rangle = \delta_{v',u'}} \langle u'|A|v'\rangle \\ &= \sum_{u'} \langle u'|A|v'\rangle \ ,\end{aligned}$$

effectively inserting and removing the identity.

3.5.6 Density Matrices

So far, when we have referred to quantum states, we meant *pure* states. Usually, preparation procedures are error prone, and one might merely know a probability distribution over states instead of the actual state. This is equivalent to having performed a measurement on a system without knowing the result of the measurement. Such a statistical mixture of states—not to be confused with a superposition of states—is represented by a *density matrix*. A density matrix ρ is defined to be a positive trace-one operator. So ρ is not only Hermitian, but also all its eigenvalues are positive and sum to one:

$$\rho = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i| \qquad \sum_i \lambda_i = 1 \qquad \lambda_i \geq 0 \ \forall i \ .$$

The eigenvalues can, therefore, be regarded as probabilities for the system to be in the state of the corresponding eigenvector.

The time evolution of density matrices derives from the unitary propagator $U(t) = e^{iHt/\hbar}$ to be

$$\rho(t) = U(t)\rho(0)U^\dagger(t) \ .$$

The probability of obtaining the result λ_i when measuring an observable A is

$$P(\lambda_i) = \mathrm{Tr}(|\varphi_i\rangle\langle\varphi_i|\rho) \ ,$$

where φ_i is the eigenvector corresponding to the eigenvalue λ_i of A.

Pure states: The density matrices with no uncertainty, *i.e.*, those with eigenvalues corresponding to a deterministic probability distribution

$$\lambda_i = \delta_{i,k} \text{ for some } k$$

are the projectors on $\mathcal{H}$.

$$\rho = |\psi\rangle\langle\psi| \text{ for some normalized } \psi \in \mathcal{H} \ ,$$

and thus in a one-to-one relation with the state vectors introduced above.

If one measures a quantum state ψ with an observable containing ψ as one of its eigenvectors, one is left deterministically with a pure state, even without registering the measurement result. The hope of measuring in the right basis to then obtain a pure state (even without knowing the result) is moot if one is left with a part of an entangled state.

Definition 2 (Separability and Entanglement)**.** States of a joint system $\mathcal{H}_A \otimes \mathcal{H}_B$ that can be written as a product,

$$\psi = \varphi_a \otimes \varphi_B \ ,$$

are called *separable*. States that are not separable are called *entangled*.

Let us consider the entangled *singlet* state

$$|\psi^-\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B\right) \ .$$

Remark 1. The state $|\psi^-\rangle$ is a *superposition* (normalized linear combination) of pure states and, hence, a pure state itself. Density matrices, on the contrary, are convex combinations of *projectors*, unless we are dealing with the special case of a pure density matrix.

Whatever the observable that Alice (or Bob, respectively) measures on her (his) part of the singlet: The probability for both results is 1/2. In particular, Alice's measurement result does not depend on measurements on Bob's side and vice versa. Thus, Alice and Bob cannot transmit messages this way—possibly faster than light. This saves us from problems with relativity. So, in summary, *parts of an entangled state are mixed states.*

To understand what Alice's part of an entangled system looks like, we introduce the *partial trace*, a linear map

$$\mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathrm{End}(\mathcal{H}_A)$$

(or $\mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathrm{End}(\mathcal{H}_B)$ respectively). The partial trace is defined by linearly extending the following map

$$\begin{aligned} \mathrm{Tr}_B : \mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_B) &\to \mathrm{End}(\mathcal{H}_A) \\ S \otimes T &\mapsto \mathrm{Tr}(T) S \ . \end{aligned}$$

The partial traces of the singlet are maximally mixed states,

$$\begin{aligned} \mathrm{Tr}_B \left(|\psi^-\rangle\langle\psi^-| \right) &= \frac{1}{2} \big(\mathrm{Tr}_B(|01\rangle\langle 01|) - \underbrace{\mathrm{Tr}_B(|01\rangle\langle 10|)}_{=0} \\ &\quad - \underbrace{\mathrm{Tr}_B(|10\rangle\langle 01|)}_{=0} + \mathrm{Tr}_B(|10\rangle\langle 10|) \big) \\ &= \frac{1}{2} \left(|0\rangle\langle 0| + |1\rangle\langle 1| \right) = \frac{\mathbb{1}_A}{2} \\ \mathrm{Tr}_A \left(|\psi^-\rangle\langle\psi^-| \right) &= \frac{\mathbb{1}_B}{2} \ . \end{aligned}$$

Importantly, the partial trace contains all the relevant information about local measurements. So, if Alice performs a measurement corresponding to an observable with eigenvector φ_i, then the probability of obtaining the result λ_i associated with this eigenvector is

$$P(\lambda_i) = \mathrm{Tr} \left((P_{\varphi_i} \otimes \mathbb{1}_B) \rho \right) \ .$$

The trace can be split into a partial trace over Bob's part and a subsequent trace over Alice's part:

$$\begin{aligned} P(\lambda_i) &= \mathrm{Tr}_A \left(\mathrm{Tr}_B \left((P_{\varphi_i} \otimes \mathbb{1}_B) \rho \right) \right) \\ &= \mathrm{Tr}_A \left(P_{\varphi_i} \, \mathrm{Tr}_B(\rho) \right) \ . \end{aligned}$$

Here we used

$$\mathrm{Tr}_A \left((A \otimes \mathbb{1}_B) \rho \right) = \mathrm{Tr}_A \left(A \, \mathrm{Tr}_B(\rho) \right) \quad \forall \ A \in \mathrm{End}(\mathcal{H}_A) \ ,$$

which follows, e.g., from expanding both sides in a product basis. The result can be extended to show that any local operations that Bob might perform on his side do not affect the probabilities of the measurement results on Alice's side. Importantly, this puts the statement that an entangled state alone does not serve to transmit information on a sound footing.

3.6 Qbits

We now apply what we have learned from quantum physics to do quantum *informatics*. Let us consider finite-dimensional systems; note that for a given dimension, there is *exactly one* Hilbert space of that dimension. The restriction to the finite-dimensional subspace — which, for instance, models the "quantum computer" in question — is admissible since quantum theory is *linear*, so the physics is the same as in the full space. The basic building block of quantum information processing as a *quantum bit* or *Qbit*, which corresponds to a Hilbert space of dimension 2. However, we see that, unlike in classical information, the understanding of single quantum bits is far from giving us an understanding of systems of two or more quantum bits: Qualitatively novel effects come into play which are essentially responsible for the power and interest of quantum informatics; or as *Ben Schumacher* put it: "It's all about entanglement."

3.6.1 One Qbit: $\mathcal{H} = \mathbb{C}^2$

A *quantum bit* or *Qbit* is represented by a physical system that has two distinguishable states, such as orthogonal polarizations of a photon, the spin of a silver atom, or a hydrogen atom with the electron being in the ground versus first excited state; these two distinguishable states are then called $|0\rangle$ and $|1\rangle$. The latter of the given examples shows that these two states of the "computational" or standard basis may be physically special (mainly in the sense that a measurement in this particular basis is simpler that in any other) even though mathematically they are not. For classical bits, the two given states would be the entire state space, whereas for quantum information, they merely form a basis. The general state of a Qbit is

$$\alpha|0\rangle + \beta|1\rangle, \text{ where } \alpha, \beta \in \mathbb{C}, \ |\alpha|^2 + |\beta|^2 = 1 \ .$$

Given the normalization condition, we still have three real degrees of freedom, which is still one too many: It is consistent with the *projector representation* of pure states that global phase factors are irrelevant and cannot be detected by any experiment: For $c \in \mathbb{C}, |c| = 1$, we have

$$P_{c|\Psi\rangle} = (c|\Psi\rangle)(c|\Psi\rangle)^* = c|\Psi\rangle\langle\Psi|\bar{c} = P_{|\Psi\rangle} \ .$$

So, roughly speaking, we have for instance $-|0\rangle = |0\rangle$. Mathematically, two vectors are said to be equivalent if and only of they differ simply by a global phase factor. Quantum states then correspond to equivalence relations with respect to that relation, sometimes called "unit rays." A representation of states which takes into account the irrelevance of global phases is the *Bloch*

sphere. In this representation of the unit rays, or projectors, "orthogonal" becomes "antipodal":

$$\begin{aligned}\alpha|0\rangle + \beta|1\rangle &= e^{i\varphi_\alpha}|\alpha||0\rangle + e^{i\varphi_\beta}|\beta||1\rangle \\ &= e^{i\varphi_\alpha}\left(|\alpha||0\rangle + e^{i(\varphi_\beta-\varphi_\alpha}\beta||1\rangle\right) \\ &= \cos\frac{\Theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\Theta}{2}|1\rangle\ .\end{aligned}$$

The state is now parametrized by two angles, $\Theta \in [0°, 180°]$ and the *relative phase* $\varphi \in [0°, 360°]$. This equals the coordinate system on Earth (including the fact that the poles do not have a well-defined east-west coordinate): The relative phase becomes a *global* one in these two cases.

The possible manipulations of a Qbit correspond to *unitary* transformations, *i.e.*, two-by-two matrices with the property that the column vectors form an orthonormal basis.

If global phase factors are irrelevant for states, then the same is true for the operations. An important example of a unitary is the *Hadamard transform*:

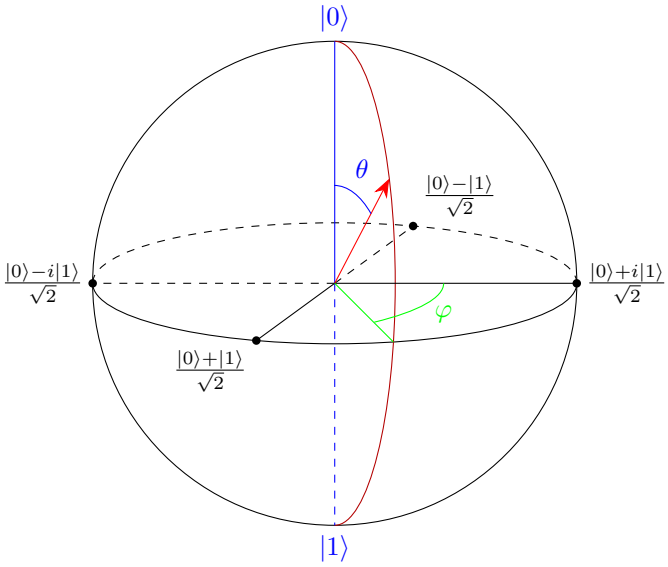
$$H := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}/\sqrt{2}\ .$$

What is the action of the Hadamard? The basis state $|0\rangle$ is mapped to $(|0\rangle + |1\rangle)\sqrt{2}$, and $|1\rangle$ to $(|0\rangle - |1\rangle)\sqrt{2}$. It looks like Hadamard's action of the Bloch sphere is a rotation around the axes connecting the circular polarizations $(|0\rangle \pm i|1\rangle)/\sqrt{2}$. However, H is not only unitary but also an *involution*, *i.e.*, $\mathrm{H}^2 = \mathbb{1}$ (and, therefore, also self-adjoint), mapping the two diagonal basis states back to the computational basis. This means that the rotation mentioned above must be combined with a second rotation, around the axis defined by the two diagonal states, by 180°. In particular, H exchanges the two *circular* basis states.

According the measurement postulate, a Qbit can be measured in any orthogonal basis. If we assume that *arbitrary* unitaries can be executed on the Qbits, then it is sufficient to be able to carry out the "standard measurement," *i.e.*, the measurement in the $\{|0\rangle, |1\rangle\}$ basis. For instance, a measurement in the diagonal basis is a Hadamard transform plus the standard measurement.

3.6.2 Two Qbits: $\mathcal{H} = \mathbb{C}^4$

In contrast to classical information, the understanding of individual Qbits is not sufficient for understanding *pairs* of Qbits, as we illustrate in this section.



States

One possible (joint) state of a two-Qbit system is that Qbit 1 is in a well-defined state, say $|\psi_1\rangle$, and Qbit 2 in a state $|\psi_2\rangle$:

$$(|\varphi_1\rangle, |\varphi_2\rangle) =: |\varphi_1\rangle \otimes |\varphi_2\rangle =: |\varphi_1\varphi_2\rangle \ .$$

Such a state of the *Qbit pair* is called a *(tensor) product state* — and obviously, the specification in the name insinuates that the product states are not all possible states of the pair: The reason for this is that the set of such products $\{|\varphi_1\varphi_2\rangle\}$ is *not* a linear space:

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} \neq |\varphi_1\rangle \otimes |\varphi_2\rangle$$

for all φ_1 and φ_2; on top of that, the state cannot be written as a mixture of such products either. This phenomenon is called *entanglement*.

The state space of a Qbit pair is the *tensor product* of the individual spaces, which is the *span* of the set of product states:

$$\mathbb{C}^2 \otimes \mathbb{C}^2 := \mathrm{span}\{\varphi_1\varphi_2\rangle\} = \mathrm{span}\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\} = \mathbb{C}^4.$$

Analogously, the space of n *Qbits* would be the n-fold tensor product

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 = \mathbb{C}^{2^n} \ ,$$

where a basis is given by the classical n-bit strings:

$$\{|i\rangle \mid i \in \{0,1\}^n\} .$$

The tensor product on the level of the vectors is carried out by the *Kronecker product*:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} .$$

By linearity, the same operation works for arbitrary pure states:

$$\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a \cdot \begin{pmatrix} c \\ d \end{pmatrix} \\ b \cdot \begin{pmatrix} c \\ d \end{pmatrix} \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix} .$$

Not of this form are entangled states, for instance, the so-called *EPR pairs, Bell states, or maximally entangled states*:

$$\begin{aligned} |\Phi^+\rangle &= \frac{1}{\sqrt{2}} (|0,0\rangle + |1,1\rangle) \\ |\Phi^-\rangle &= \frac{1}{\sqrt{2}} (|0,0\rangle - |1,1\rangle) \\ |\Psi^+\rangle &= \frac{1}{\sqrt{2}} (|0,1\rangle + |1,0\rangle) \\ |\Psi^-\rangle &= \frac{1}{\sqrt{2}} (|0,1\rangle - |1,0\rangle) \end{aligned}$$

Scalar product

The scalar product is fully defined by the fact that the pairwise tensor product of states of two respective orthogonal bases of the individual spaces is again an orthonormal basis of the composed space. What does the so-defined scalar product on tensor-product states look like? Let

$$\psi_0 := \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} , \ \varphi_0 := \begin{pmatrix} \gamma_0 \\ \delta_0 \end{pmatrix} , \psi_1 := \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} , \ \varphi_1 := \begin{pmatrix} \gamma_1 \\ \delta_1 \end{pmatrix} .$$

Then

$$\begin{aligned}(|\psi_0\rangle \otimes |\varphi_0\rangle\,,\, |\psi_1\rangle \otimes |\varphi_1\rangle) &= (|\psi_0\rangle^\dagger \otimes |\varphi_0\rangle^\dagger)\ (|\psi_1\rangle \otimes |\varphi_1\rangle) \\ &= (\bar{\alpha_0}\bar{\gamma_0}\,,\, \bar{\alpha_0}\bar{\delta_0}\,,\, \bar{\beta_0}\bar{\gamma_0}\,,\, \bar{\beta_0}\bar{\delta_0}) \cdot \begin{pmatrix} \alpha_1\gamma_1 \\ \alpha_1\delta_1 \\ \beta_1\gamma_1 \\ \beta_1\delta_1 \end{pmatrix} \\ &= \bar{\alpha_0}\alpha_1\bar{\gamma_0}\gamma_1 + \bar{\alpha_0}\alpha_1\bar{\delta_0}\delta_1 + \bar{\beta_0}\beta_1\bar{\gamma_0}\gamma_1 + \bar{\beta_0}\beta_1\bar{\delta_0}\delta_1 \\ &= (\bar{\alpha_0}\alpha_1 + \bar{\beta_0}\beta_1)(\bar{\gamma_0}\gamma_1 + \bar{\delta_0}\delta_1) \\ &= \langle\psi_0|\psi_1\rangle \cdot \langle\varphi_0|\varphi_1\rangle\ .\end{aligned}$$

In short:

$$\langle\psi_0\varphi_0|\psi_1\varphi_1\rangle = \langle\psi_0|\psi_1\rangle \cdot \langle\varphi_0|\varphi_1\rangle\ .$$

In particular, two tensor products are orthogonal if and only if the corresponding components are orthogonal in at least one of the partial spaces.

Operations

An example of an operation that can be carried out on a pair of Qbits is a product operation. Let us define it on the set of product states, which is a generating system of the full space; the operation is thus uniquely defined by linearity:

$$U \otimes V(|\psi\rangle \otimes |\varphi\rangle) := U|\psi\rangle \otimes V|\varphi\rangle\ .$$

This map is unitary as long as both U and V are: It is easy to see that orthogonal tensor products are mapped to orthogonal tensor product along our understanding of orthogonality above.

Example 3.6.1. The matrix Hadamard on two Qbits can again be computed through the Kronecker product:

$$\mathrm{H}\otimes\mathrm{H} = \frac{1}{2}\left[\begin{array}{c|c} 1\cdot\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} & 1\cdot\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ \hline 1\cdot\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} & -1\cdot\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \end{array}\right] = \frac{1}{2}\left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ \hline 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{array}\right]$$

This representation is not very intuitive or transparent. A better understanding can be obtained from this formula: For $b \in \{0, 1\}$,

$$\mathrm{H}\,|b\rangle = \frac{|0\rangle + (-1)^b|1\rangle}{\sqrt{2}}\ .$$

Accordingly, for two Qbits,

$$\begin{aligned}\mathrm{H}^{\otimes 2}\,|b_1b_2\rangle &= \mathrm{H}\,|b_1\rangle \otimes \mathrm{H}\,|b_2\rangle \\ &= \frac{1}{2}\big((|0\rangle + (-1)^{b_1}|1\rangle) \otimes (|0\rangle + (-1)^{b_2}|1\rangle)\big) \\ &= \frac{1}{2}\big(|00\rangle + (-1)^{b_2}|01\rangle + (-1)^{b_1}|10\rangle + (-1)^{b_1\oplus b_2}|11\rangle\big)\ .\end{aligned}$$

For n Qbits:

$$\mathrm{H}^{\otimes n}\,|b\rangle = \frac{1}{2^{n/2}} \sum_{i\in\{0,1\}^2} (-1)^{b\cdot i}|i\rangle\ ,$$

where

$$b \cdot i := \bigoplus_{j=1}^{n} b_j \wedge i_j\ .$$

In particular,

$$\mathrm{H}^{\otimes n}\,|b\rangle = \frac{1}{2^{n/2}} \sum_{i\in\{0,1\}^2} |i\rangle$$

will be the initial state for most quantum algorithms, and as a equal superposition of all classical inputs, the basis for "Quantum parallelism."

Measurements

Also for two Qbits, we have that general transformations plus one standard measurement allows for *general* measurements. (So far, we have only considered product operations, we return to general ones later.)

It is, however, also possible to measure *only one* of a pair of Qbits. Let us assume a general state

$$|\Psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle\ .$$

What if only the first Qbit is measured? What are the statistics? What is the state of the second Qbit (conditioned on the two measurement outcomes)?

The first question is simpler: The probability of 0 is $p_0 := |\alpha|^2 + |\beta|^2$ and of 1, $p_1 = 1 - p_0 = |\gamma|^2 + |\delta|^2$.

In order to answer the second question, we rewrite the state as:

$$\begin{aligned}|\Psi\rangle &= |0\rangle \otimes (\alpha|0\rangle + \beta|1\rangle) + |1\rangle(\gamma|0\rangle + \delta|1\rangle) \\ &= \sqrt{p_0}|0\rangle \otimes \left(\frac{\alpha}{\sqrt{p_0}}|0\rangle + \frac{\beta}{\sqrt{p_0}}|1\rangle\right) + \sqrt{p_1}|1\rangle \otimes \left(\frac{\gamma}{\sqrt{p_1}}|0\rangle + \frac{\delta}{\sqrt{p_1}}|1\rangle\right) \\ &= \sqrt{p_0}|0\rangle \otimes |\psi_0\rangle + \sqrt{p_1}|1\rangle \otimes |\psi_1\rangle\ .\end{aligned}$$

Then, $|\psi_0\rangle$ and $|\psi_1\rangle$ is the state of the second Qbit, given the measurement of the first yields 0 and 1, respectively. Since the state of Qbit 2 does not depend on what happens to the first, it is, even if the first Qbit is not measured at all:

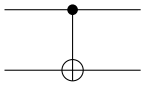
$$p_0 \cdot P_{|\psi_0\rangle} + p_1 \cdot P_{|\psi_1\rangle} \ .$$

This is called the *partial trace* of $|\Psi\rangle$ if the first Qbit is "traced out." We, again, see what we have already observed: *Parts of entangled states are mixtures.*

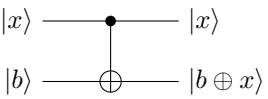
3.6.3 Back to Operations: The CNOT Gate

What holds for states is just as true for the operations on pairs of Qbits: The products are not the whole story. Indeed, there exist unitaries of pairs of Qbits which cannot be decomposed into two individual operations. Remarkably, it is enough to have *one single two-Qbit operation* — the CNOT we discuss here — plus *unary (one-bit unitary) operations*, and *every unitary on an arbitrary number of Qbits* becomes possible.

The CNOT can be seen as a made-reversible XOR or, as the name says, a "controlled-NOT" gate.



The action of the CNOT on the classical basis of the space of Qbit pairs is as follows:



(Note again here that the Latin letters b and x are supposed to take only the values 0 and 1; in contrast, Greek letters stand for arbitrary, generally *non*-classical, quantum states.) In matrix form, the CNOT looks as follows.

$$\begin{array}{c} \\ 00 \\ 01 \\ 10 \\ 11 \end{array} \begin{array}{c} \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & & 1 \\ & & 1 & \end{pmatrix} \end{array}$$

What have we won now? After all, we are able to compute the XOR also classically. In fact, although the action of the gate is now uniquely determined through linearity, there are still surprises waiting for us. Let us, for instance, look at what happens if we enter *diagonal* states into the gate. The input is then, in the standard basis,

$$\frac{1}{2}[|00\rangle - |01\rangle + |10\rangle - |11\rangle] \ .$$

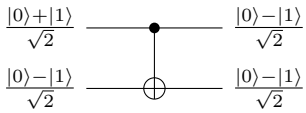
The CNOT maps it to the output

$$\frac{1}{2}[|00\rangle - |01\rangle + |11\rangle - |10\rangle] \ ,$$

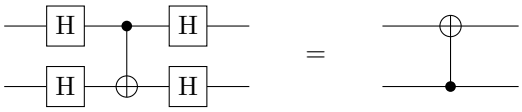
which is a product state again:

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \ .$$

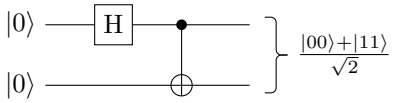
It is still true that the output is not entangled. But the roles of the wires, *i.e.*, source and control, have been swapped.



Indeed, it is an easy exercise to show that the CNOT acting on *diagonal* states is again a CNOT — with swapped roles, however:



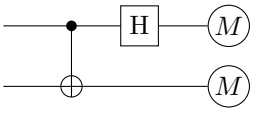
More interestingly even, the CNOT can *generate entanglement*, *i.e.*, map products into non-products:



In fact, the combination of a Hadamard on the control wire and the CNOT is a basis change between the standard basis and the "Bell basis"

$$\{|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle\} \ .$$

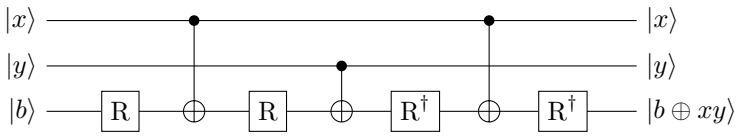
In particular,



is a *Bell measurement*, i.e., a measurement in the Bell basis.

As mentioned, the CNOT gate is universal. We show here only its *classical* universality: It allows for computing arbitrary classical functions. For this,

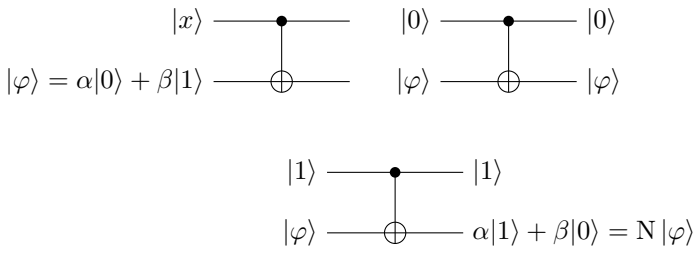
it is enough to show that it allows, together with single-Qbit operations, to obtain a *Toffoli*. This is done with the following circuit:



Here, R is a rotation around 22.5°,

$$\begin{bmatrix} \cos 22.5^\circ & -\sin 22.5^\circ \\ \sin 22.5^\circ & \cos 22.5^\circ \end{bmatrix}.$$

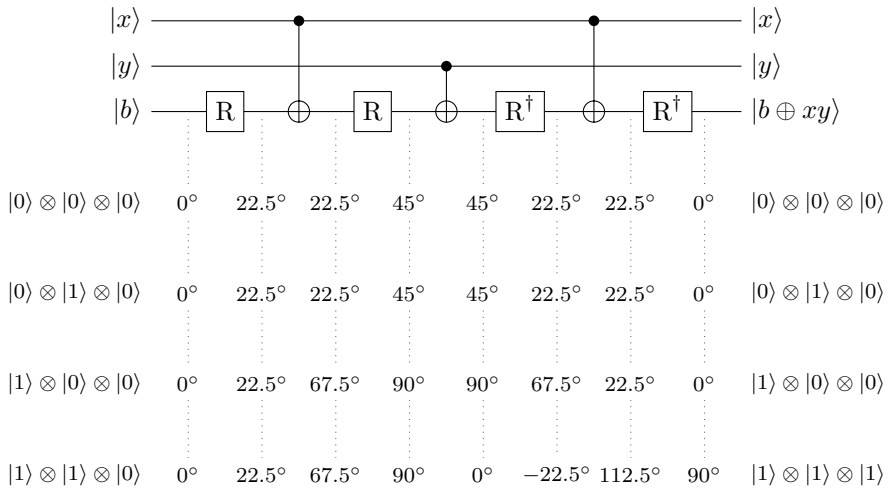
For the proof, let us first understand the action of the CNOT if the control $|x\rangle \in \{|0\rangle, |1\rangle\}$ is classical and the source $|\varphi\rangle = \alpha|0\rangle + \beta|1\rangle$ arbitrary:



If $x = 0$, the gate is the identity. If, on the other hand, $x = 1$, then the output is still a product, and the control is unchanged, but $\varphi = \alpha|0\rangle + \beta|1\rangle$ is mapped to $\alpha|1\rangle + \beta|0\rangle$. More precisely, the negation

$$N = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

is executed to the second Qbit, which geometrically corresponds to a reflection around the (positive) diagonal.

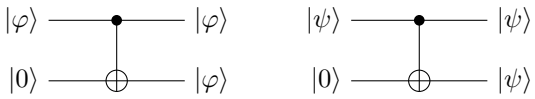


The case $b = 1$ follows accordingly from the fact that the circuit is unitary and $|xy1\rangle$ is orthogonal to $|xy0\rangle$.

3.6.4 Cloning, Pseudo-Cloning, and Pseudo-Measurements

It is not difficult to copy classical bits. For instance, the CNOT gate for classical inputs does it if the source is 0.

It is, therefore, natural to ask whether the CNOT, with $|0\rangle$ as source, also allows for "cloning" — another word for copying used in that context — *quantum* information:



The unitarity of the CNOT then implies

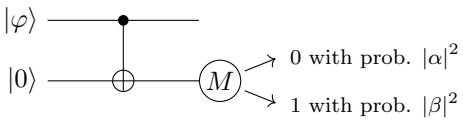
$$\langle\psi|\varphi\rangle\langle 0|0\rangle = \langle\psi|\varphi\rangle\langle\psi|\varphi\rangle \ ,$$

which can hold only if

$$\langle\psi|\varphi\rangle \in \{0, 1\} \ .$$

Only *parallel or orthogonal* states can be cloned by the CNOT. Note that all we have used was the *unitarity* property of the gate. In other words, the cloning operation is not unitary and, hence, not allowed by quantum theory: This is the *no-cloning theorem*.

If the CNOT does not clone quantum states, then what *does* it do? (In fact, we have already seen part of the answer: It generates entanglement.)



If the second Qbit is measured in the standard basis, then the statistics are, actually, the same *as if the input state actually had been cloned*. This is not in contraction to the no-cloning theorem, since it holds only for the classical basis. The same is true for the first Qbit: Under standard measurements, the statistics are as if it were the original state since its state is the mixture

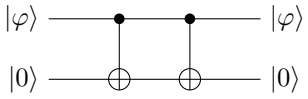
$$|\alpha|^2 P_{|0\rangle} + |\beta|^2 P_{|1\rangle} \ .$$

In fact, the measurement outcomes on both wires would be always identical since the joint state is

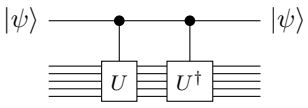
$$\alpha|00\rangle + \beta|11\rangle \ .$$

Because of this, this action of the CNOT is sometimes called *"pseudo-cloning."*

What if the measurement on the second wire is not done? Obviously, this should not change the state on the first wire. This means that the CNOT leads to the same transition of the first Qbit as if a measurement (on this first Qbit) had happened. Because of this, the CNOT can also be seen as a "pseudo-measurement." It is not a measurement since there is no outcome, and because it is *reversible*: The CNOT is an involution, *i.e.*, self-inverse.



Another way to see the measurement process is by replacing the second Qbit with a very high-dimensional system, and the CNOT by some controlled unitary action on the large system containing the measurement apparatus, the physicist in the lab, the laboratory, the environment...



In this view, a measurement would always be *in principle* reversible. Such a measurement does not induce a "collapse," but *decoherence*.

The same picture, finally, also illustrates *disturbance*: If a quantum computation interacts with only one single binary degree of freedom not belonging to the computation, for instance, since one of the Qbits of the computation controls a CNOT to the "vagabond," the state of the computer turns from a superposition into a mixture. If that, *e.g.*, air molecule or photon escapes, then the process is irreversible and the computation fails. (This, again, was the reason for our detailed study of how to "uncompute" junk in an "orderly" way.)

Chapter 4

Quantum Communication

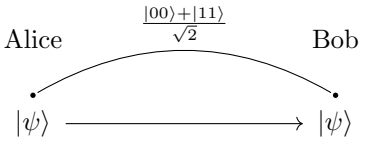
4.1 Teleportation

Quantum teleportation, proposed by Bennett, Brassard, Crépeau, Josza, Peres, and Wootters [3], is certainly one of the most exciting, and inspiring results in the field. (The story goes that one of the inventors preferred the term "teleferism" for avoiding the, in his eyes, ugly Greek-Latin mixture.)

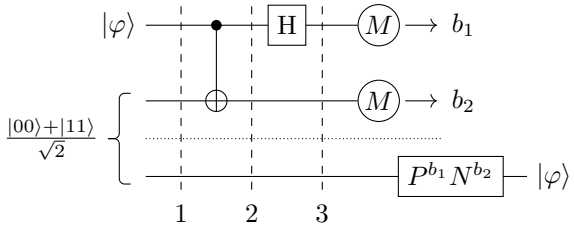
What is teleportation? An object is teleported from A to B if it is first at A, at the end at B, but has never been anywhere in-between. The transfer does not have to be instantaneous, which would, by the way, also contradict relativity.

In the end, teleportation allows for carrying over the ability to transmit *quantum* information between two parties to a later point in time, when only a classical channel is available. This is comparable to the scenario of traditional, secret-key cryptography: The availability of a confidential channel at an earlier point in time allows for transforming an insecure channel available later into a secure one. A key is exchanged over the secure channel which is then, later, used for encryption.

A priori, sending an unknown quantum state over a classical state is impossible. If, however, the parties additionally share entanglement, then the task becomes possible:



The teleportation circuit is the following:



The states in the respective positions are the following. At Position 1, we have

$$|\psi\rangle \otimes \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{\alpha}{\sqrt{2}}|000\rangle + \frac{\alpha}{\sqrt{2}}|011\rangle + \frac{\beta}{\sqrt{2}}|100\rangle + \frac{\beta}{\sqrt{2}}|111\rangle \ .$$

After the application of the CNOT two the first and the second Qbit, we have at Position 2:

$$\begin{aligned} &\frac{\alpha}{\sqrt{2}}|000\rangle + \frac{\alpha}{\sqrt{2}}|011\rangle + \frac{\beta}{\sqrt{2}}|110\rangle + \frac{\beta}{\sqrt{2}}|101\rangle \\ &= \alpha|0\rangle \otimes \frac{|00\rangle + |11\rangle}{\sqrt{2}} + \beta|1\rangle \otimes \frac{|10\rangle + |01\rangle}{\sqrt{2}} \ . \end{aligned}$$

This is a three-entangled state. After the Hadamard gate on the first Qbit, the joint state becomes at Position 3

$$\begin{aligned} &\alpha \ \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|00\rangle + |11\rangle}{\sqrt{2}} + \beta \ \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \frac{|10\rangle + |01\rangle}{\sqrt{2}} \\ &= \frac{\alpha}{2}\big(|000\rangle + |011\rangle + |100\rangle + |111\rangle\big) + \frac{\beta}{2}\big(|010\rangle + |001\rangle - |110\rangle - |101\rangle\big) \\ &= \frac{1}{2}\Big(|00\rangle \otimes \underbrace{(\alpha|0\rangle + \beta|1\rangle)}_{=|\Psi\rangle} + |01\rangle \otimes \underbrace{(\alpha|1\rangle + \beta|0\rangle)}_{=N|\Psi\rangle} + \\ &\qquad |10\rangle \otimes \underbrace{(\alpha|0\rangle - \beta|1\rangle)}_{=|P\Psi\rangle} + |11\rangle \otimes \underbrace{(\alpha|1\rangle - \beta|0\rangle)}_{=NP|\Psi\rangle}\Big) \ . \end{aligned}$$

Here, N and P stand for the *negation* and *conditional phase flip*, respectively:

$$N := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \ , \quad P := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \ .$$

Now, Alice measures the state and sends the result $\{b_1, b_2\} \in \{0,1\}^2$ to Bob who, upon reception, applies the transform $P^{b_1}N^{b_2}$ in order to end up with the state $|\Psi\rangle$ in all four cases.

The main application of teleportation are *quantum repeaters*, allowing for so-called *entanglement swapping*: Cryptographic protocols exist that are device-independent, *i.e.*, where no trust in the manufacturer and actually not even in the correctness of quantum theory is required. The protocols are based on the parties sharing maximal entangled states. Now, the problem is that it is not an easy task to establish such entanglement over large distances. The further particles are transported, the likelier it is that they interact with their environment, and, consequently, that they lose their initial entanglement. Teleportation, however, can help. Imagine that Alice and Bob are too far apart to directly exchange entanglement, but that Charlie is in the middle to establish a singlet with both Alice and Bob. Charlie can then act as a *quantum repeater*, using the EPR pair shared with Bob for teleporting his share of the Bell state with Alice to Bob. In fact, it is a property of teleportation that, if part of an entangled state is sent, then this entanglement is preserved. What Charlie really has to do is actually a Bell measurement on his pair and send its result either to Alice or to Bob. A drawback of this in the cryptographic context is that the inner node Charlie has to be trusted.[1]

Does teleportation not contradict relativity? At first sight, the answer is clearly no: After all, Alice must transmit a classical message — which cannot arrive faster than at the speed of light. Still, it is an interesting observation in this context that, with probability 1/4, Bob does not have to do anything, but already *has* the state $|\Psi\rangle$ on his wire. Since nothing has happened to Bob's Qbit, this means that there has always been $|\Psi\rangle$ on his wire, in that case, even before his reception of Alice's message — maybe even before Ψ even existed? Is that not problematic? Maybe not: After all, I can flip a coin now, and with a probability of 1/2, it will correctly indicate the weather in exactly 10 years; nothing weird about that. But then, a next thought might be that the state space of a quantum system is not finite (as with the coin, where it is of size 2), so is it not problematic to have a probability of 1/4 to have Ψ on that wire before Alice's message arrives? Does it not mean that the corresponding mixed state on Bob's wire is somehow a function of $|\Psi\rangle$? In fact, *no*: We invite the reader to verify that

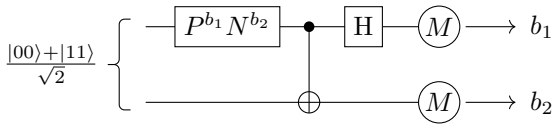
$$\frac{1}{4}\left(P_{|\Psi\rangle} + P_{N|\Psi\rangle} + P_{P|\Psi\rangle} + P_{NP|\Psi\rangle}\right) = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} .$$

Again, the density-matrix formalism and the corresponding postulate saves us from serious trouble with relativity.

[1] Even in the face of this, the Chinese government has not been shy to invest billions for realizing such a "quantum backbone 'secure' network."

4.2 Superdense Coding

The scenario of *superdense coding* is exactly the inverse of that of teleportation: Given a Qbit channel between Alice and Bob, can they transmit *classical* information? Of course, the first answer is *yes*: With each Qbit sent, one classical bit (Cbit) can be sent. *Alexander Holevo* [11] has proven, however, that this is optimal. (This fact is not so surprising since measuring a Qbit yields only one classical bit of information.) More classical information, however, can travel with a single Qbit *if the parties additionally share entanglement*. This comes as a surprise, since entanglement *alone* does not allow for transmitting *any* information. This "superadditivity" of resources is sometimes called "activation."



Assume Alice wants to transmit a pair of bits $\{b_1, b_2\}$ to Bob. She applies the transformation

$$P^{b_1}N^{b_2}$$

to her Qbits, *i.e.*, to her half of the Bell state. Note that this is exactly the same transformation as the one Bob applies in teleportation in function of the two bits he receives from Alice. She then sends this Qbit to Bob, who performs a *Bell measurement* on his two Qbits — which tells him the pair of bits Alice wanted to send. Interestingly, despite the quite opposite goals pursued in teleportation and superdense coding, the transformations are actually the same, only carried out in reverse order. A reading is that one of the two classical bits travels along the transmitted Qbit, whereas the other travels back in time with Alice's half of the EPR pair, to the point where the pair was generated and then "back to the future" to Bob along the other half of the pair. This may sounds adventurous, but quantum mechanics certainly *does* challenge how we usually think about causality — at least since *John Bell*'s disturbing discovery.

Chapter 5

Simple Algorithms

5.1 n Qbits

The state space of systems of n Qbits is the n-fold tensor product of $\mathbb{C}^2$ with itself:

$$\mathcal{H} = \underbrace{\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{n\text{times}} \equiv \mathbb{C}^{2^n} \ .$$

A basis of the space is given by the classical basis characterized through the set of all (classical) n-bit strings,

$$\{|\mathbf{i}\rangle \,|\, \mathbf{i} \in \{0,1\}^n\} \ .$$

The action of the n-fold, bitwise *Hadamard transform* is

$$\begin{aligned} \mathrm{H}^{\otimes n} |\mathbf{i}\rangle &= \bigotimes_{l=1}^{n} \left(\frac{|0\rangle + (-1)^{i_l} |1\rangle}{\sqrt{2}} \right) \\ &= \frac{1}{2^{n/2}} \sum_{\mathbf{j} \in \{0,1\}^n} (-1)^{\mathbf{i}\cdot\mathbf{j}} |\mathbf{j}\rangle \ , \end{aligned}$$

where the "logical scalar product" $\mathbf{i} \cdot \mathbf{j}$ is defined as

$$\mathbf{i} \cdot \mathbf{j} := \bigoplus_{k=1}^{n} i_k \wedge j_k \ .$$

Note that this scalar product expresses the parity of the number of bit positions, from 1 to n where both involved strings have a 1. In particular, the product is always 0 if *one* of the two strings is the all-zero string $\mathbf{0}$. This means that the Hadamard applied to this vector has only + signs: It is the

equal superposition of all classical states, and it will be the input for all quantum algorithms, realizing "quantum parallelism." Note also that the state $|\mathbf{0}\rangle$ as a part of the Hadamard transform applied to a general state also always has a positive sign, *i.e.*, probability amplitude.

5.2 The Secret Mask

The first algorithm we discuss is — despite not being historically the first — the one due to *Ethan Bernstein* and *Umesh Vazirani* [4].

Let $\mathbf{s} \in \{0,1\}^n$. Then $f_\mathbf{s}$ is the following function from n bits to one bit:

$$f_\mathbf{s} := \mathbf{s} \cdot \mathbf{x} = \bigoplus_{i=1}^{n} (s_i \wedge x_i) \ .$$

Let us imagine a black box implementing this function f:

$$\mathbf{x} \in \{0,1\}^n \longrightarrow \boxed{\mathbf{s} \cdot \mathbf{x}} \longrightarrow f_s(x) \ .$$

To find out $\mathbf{s}$, one has to make (exactly) n queries:

$$\text{Input} e_i = 00 \ldots 0100 \ldots 0 \longrightarrow s_i \ .$$

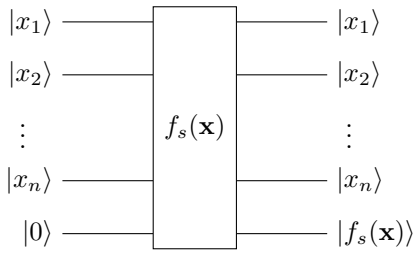
This is optimal: Only one bit of information per query is learnt since only one *physical* bit is outputted. What if the string is hidden in a *quantum* circuit?

Note first that the above classical circuit cannot be *directly* translated to a quantum circuit since it is not reversible. The first step is, hence, to express it as a reversible function and circuit: The function $\tilde{f}_\mathbf{s}$ maps $n+1$ bits to $n+1$ bits,

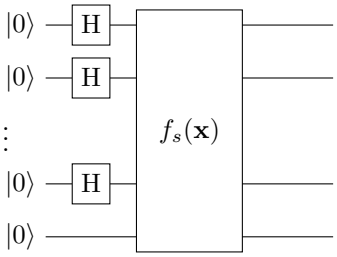
$$\tilde{f}_\mathbf{s}(\mathbf{x}, b) = (\mathbf{x}, b \oplus f_\mathbf{s}(\mathbf{x})) \ .$$

The lower bound derived above still applies to this circuit, since only one of the $n+1$ output bits is "informative" — the others are already part of the input.

The reversible circuit can now be directly interpreted as a quantum circuit.



In particular, it can be queried with the equal superposition of all classical inputs that can be obtained by applying the Hadamard to each of the n input wires and by inputting $|0\rangle$ on the last, the output wire.

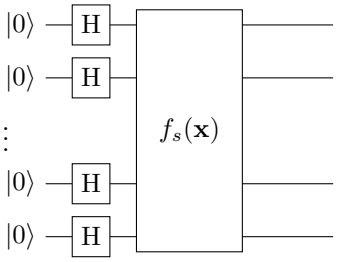


Indeed, the joint state of the output wires is then

$$\frac{1}{2^{n/2}} \sum_{\mathbf{j}} (|\mathbf{j}\rangle \otimes |f_{\mathbf{s}}(\mathbf{j})\rangle) \ .$$

This is *quantum parallelism at its best*: A single execution of the circuit allows for generating a state containing the function value for all inputs at the same time. However, it is also true that such parallelism *alone* is not very helpful. Indeed, when the resulting state is measured in the standard basis, then one obtains a random input together with the corresponding output. This is similar to calling the classical circuit once for a randomly chosen input.

Can it help to use a nonclassical input also on the last wire? Let us provide the state $\mathrm{H}\,|0\rangle$ on that wire as well:



The input state

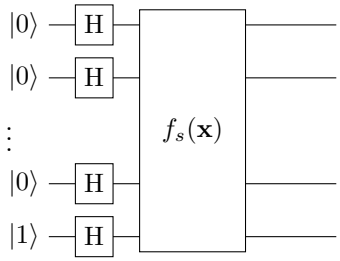
$$\frac{1}{2^{n/2}} \sum_{\mathbf{j}} |\mathbf{j}\rangle \oplus \frac{|0\rangle + |1\rangle}{\sqrt{2}} = \frac{1}{2^{n/2}} \sum_{\mathbf{j}} \frac{1}{\sqrt{2}} (|\mathbf{j}\rangle|0\rangle + |\mathbf{j}\rangle|1\rangle)$$

is then mapped to

$$\frac{1}{2^{n/2}} \sum_{\mathbf{j}} \frac{1}{\sqrt{2}} \big(|\mathbf{j}\rangle|f_{\mathbf{s}}(\mathbf{j})\rangle + |\mathbf{j}\rangle|f_{\mathbf{s}}(\mathbf{j} \oplus 1)\rangle\big) \ .$$

This is the same state as the input: The circuits acts as the identity and cannot be of any help. The reason is that addition is commutative, and all that happens is that the order of the two terms in the sum is swapped.

Now, a slight modification helps: If the input state to the last wire is $\mathrm{H}\,|1\rangle$ instead of $\mathrm{H}\,|0\rangle$, then the addition becomes a subtraction, and this operation is not commutative:



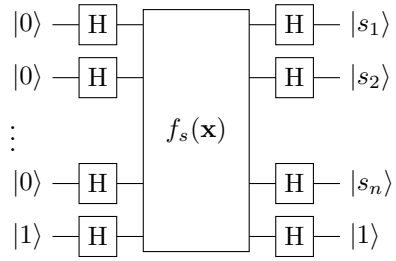
The output state becomes

$$\frac{1}{2^{n/2}}\sum_{\mathbf{j}}\frac{1}{\sqrt{2}}\big(|\mathbf{j}\rangle|f_{\mathbf{s}}(\mathbf{j})\rangle - |\mathbf{j}\rangle|f_{\mathbf{s}}(\mathbf{j}\oplus 1)\rangle\big) = \frac{1}{2^{n/2}}\sum_{\mathbf{j}}(-1)^{f_{\mathbf{s}}(\mathbf{j})}\cdot\underbrace{\frac{|\mathbf{j}\rangle|0\rangle - |\mathbf{j}\rangle|1\rangle}{\sqrt{2}}}_{|\mathbf{j}\rangle\otimes\frac{|0\rangle-|1\rangle}{\sqrt{2}}}\ .$$

The state on the last (result) wire is constant, $\mathrm{H}\,|1\rangle$, whereas on the first n wires, we have

$$\frac{1}{2^{n/2}}\sum_{\mathbf{j}}(-1)^{f_{\mathbf{s}}(\mathbf{j})}|\mathbf{j}\rangle\ .$$

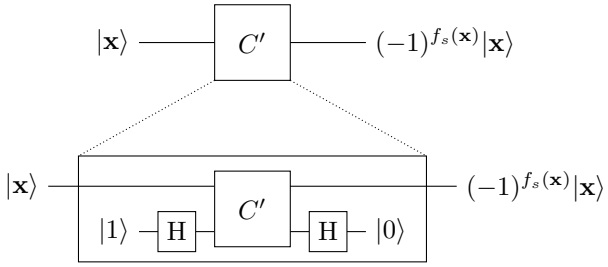
The final observation is that, for the given function, the sign factor equals $(-1)^{\mathbf{j}\cdot\mathbf{s}}$, and the n Qbits are in the state $\mathrm{H}^{\otimes n}\,|\mathbf{s}\rangle$: Thus, applying H again and then measuring yields (all n bits of) $|\mathbf{s}\rangle$ on the first n Qbits:



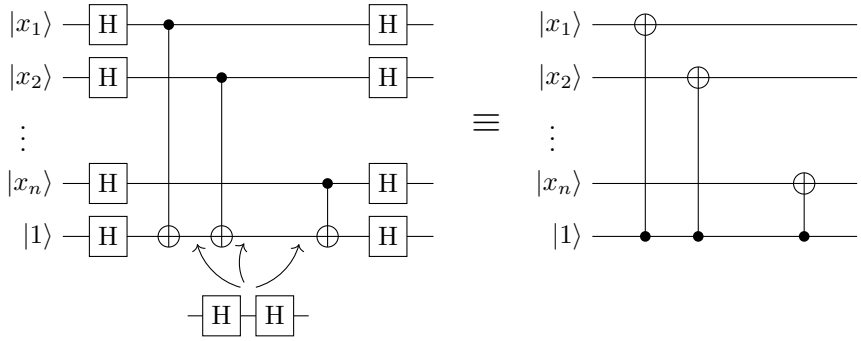
The problem can be completely solved with *one single call* to the quantum circuit.

The algorithm combines two tricks: Quantum parallelism and *phase kick-back*: When the state put on the result wire is $\mathrm{H}\,|1\rangle$, then the function's output

is encoded in a phase factor of the state of the input wires, for classical inputs: The resulting circuit maps $|x\rangle$ to $(-1)^{f(x)}|x\rangle$.



A simplified view of the algorithm was proposed by *David Mermin* [13]. Note first that the CNOT gate, combined with *Hadamards* on all four wires, is a CNOT with swapped roles. Expressed in CNOTs, the circuit implementing the function in question looks as follows: If the bit $s_1 = 1$, then there is a CNOT gate from the i-th input wire to the output Qbit; if $s_1 = 0$, there is no such gate. Using the Hadamards at the entry and exit, as well as added pairs of Hadamards in-between the CNOTs, leads to the following simple view of the problem:



The CNOTs being inverted illustrates the effect of phase kickback in this example: The full value of $\mathbf{s}$ is written to the n input wires and can be directly measured.

5.3 The Deutsch/Josza Algorithm

David Deutsch is recognized as the inventor of quantum computing by proposing a weaker variant of the following algorithm [5] for the special case $n = 1$. Let

$$f : \{0,1\}^n \to \{0,1\} ,$$

together with the promise that f is either *constant* ($f \equiv 0$ or $f \equiv 1$) or *balanced*, *i.e.*,

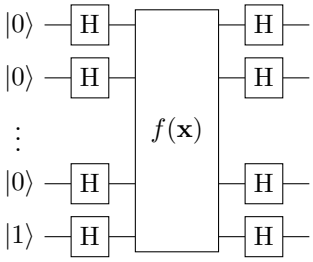
$$|\{x|f(x)=0\}| = |\{x|f(x)=1\}| = 2^{n-1} \ .$$

The problem is how to decide which of the two is the case. If the function is given as a classical circuit, then the worst-case, zero-error number of calls required to solve the problem is

$$\frac{2^n}{2} + 1 \ .$$

The quantum algorithm offers an exponential advantage: One single call is sufficient.

Let us directly combine the same two tricks



The state after the first Hadamard gates is then

$$\frac{1}{2^{n/2}} \sum_{\mathbf{j}} |\mathbf{j}\rangle \otimes \mathrm{H}\,|1\rangle \ .$$

The state after the f-gate is then, via phase kickback,

$$\frac{1}{2^{n/2}} \sum_{\mathbf{j}} (-1)^{f(\mathbf{j})} |\mathbf{j}\rangle \otimes \mathrm{H}\,|1\rangle \ .$$

Let us sort these terms with respect to their sign, representing only the first n Qbits:

$$\sum_{\mathbf{j}:f(\mathbf{j})=0} \frac{1}{2^{n/2}} |\mathbf{j}\rangle - \sum_{\mathbf{j}:f(\mathbf{j})=1} \frac{1}{2^{n/2}} |\mathbf{j}\rangle \ .$$

The final n-fold *Hadamard* is applied to this state. The resulting state would then be represented by a double sum and looks complicated, in particular due to the signs. The probability amplitude of only the output state $|000\cdots0\rangle$ would, however, be simpler to compute since the corresponding sign is always $+$: It is equal to

$$\sum_{\mathbf{j}:f(\mathbf{j})=0} \frac{1}{2^{n/2}} \cdot \frac{1}{2^{n/2}} - \sum_{\mathbf{j}:f(\mathbf{j})=1} \frac{1}{2^{n/2}} \cdot \frac{1}{2^{n/2}} = \sum_{\mathbf{j}:f(\mathbf{j})=0} \frac{1}{2^n} - \sum_{\mathbf{j}:f(\mathbf{j})=1} \frac{1}{2^n} \ .$$

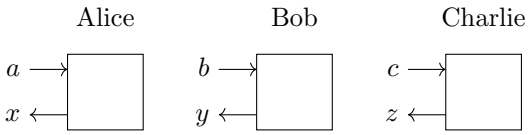
This is equal to 0 if f is balanced, to 1 if $f \equiv 0$, and to -1 if $f \equiv 1$: This means that the output $000\cdots0$ is *always* measured when f is constant, but *never* when the function is balanced: The question is answered. (Interestingly, a second call is necessary to determine which constant value the function assumes in the first case.)

Chapter 6

Intermezzo: "Pseudo-Telepathy"

One of the weirdest phenomena of quantum theory are *non-local correlations*, *i.e.*, correlations stronger than what is explainable by pre-shared classical information. Sometimes, the phenomenon is described in terms of a *game* that can be won with higher probability using quantum states than without. In the special case where quantum theory allows to win such a game *with certainty* whereas classical strategies do not, the phenomenon has been called "pseudo-telepathy": It looks as if two parties would communicate for winning the game, but they do not. Here, "winning the game" means to achieve a certain correlation in the answers the party give when asked respective question. The parties may agree on a strategy beforehand, but are not allowed to communicate anymore after having received their respective questions.

The following example of a pseudo-telepathy game stems from *David Mermin* [12]. It involves three parties, receiving questions $a, b, c \in \{0, 1\}$, respectively, and responding with $x, y, z \in \{0, 1\}$:



The game is won if the following condition is satisfied:

$$\text{If } a \oplus b \oplus c = 1, \text{ then } x \oplus y \oplus z = a \wedge b \wedge c \ .$$

(The condition $a \oplus b \oplus c = 1$ is often assumed to be satisfied and called a *promise*.)

Can this game be won with certainty with a strategy in the form of classical information? We can without loss of generality assume such a strategy to be deterministic, and, thus, determine any party's output given all possible inputs to that party. Explicitly, the strategy is a sextet of bits $x_0, x_1, y_0, y_1, z_0, z_1$, where, for instance, x_0 is the first party's output given her input is $a = 0$, *etc.* The game is *always won* if these bits, i.e., binary variables, satisfy the following conditions for the four cases in question:

$$\begin{aligned} x_0 \oplus y_0 \oplus z_1 &= 0 \\ x_0 \oplus y_1 \oplus z_0 &= 0 \\ x_1 \oplus y_0 \oplus z_0 &= 0 \\ x_1 \oplus y_1 \oplus z_1 &= 1 \ . \end{aligned}$$

When the left and right sides of the equations are $\oplus$-summed up, we get, $0 = 1$. A classical strategy that allows one to always win the game cannot exist.

What if the parties share not only classical information, but also quantum entanglement? Let us consider the following state:

$$\begin{aligned} |\Psi\rangle &= \frac{1}{2\sqrt{2}} \sum_{i,j,k\in\{0,1\}} (-1)^{\mathrm{maj}(i,j,k)} |i,j,k\rangle \\ &= \frac{1}{2\sqrt{2}} \big(|000\rangle + |001\rangle + |010\rangle - |011\rangle \\ &\qquad + |100\rangle - |101\rangle - |110\rangle - |111\rangle \big) \ . \end{aligned}$$

A first observation is that the state is symmetric between the three parties. When they all directly measure their respective Qbits in the standard basis, the phases disappear, and all they get are three independent coin flips — certainly not something helpful for establishing a "magic" correlation. What if one of the parties applies a *Hadamard* before measuring? In order to answer this, let us first factor the first Qbit away from the second and the third:

$$\begin{aligned} |\Psi\rangle = \frac{1}{2} \Bigg(& \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |00\rangle + \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes |01\rangle \\ & + \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes |10\rangle - \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |11\rangle \Bigg) \ . \end{aligned}$$

Now, the *Hadamard* can easily be applied to the first Qbit:

$$(\mathrm{H} \otimes \mathbf{1} \otimes \mathbf{1}) |\Psi\rangle = \frac{1}{2} \left(|0\rangle \otimes |00\rangle + |1\rangle \otimes |01\rangle + |1\rangle \otimes |10\rangle - |0\rangle \otimes |11\rangle \right) \ .$$

If the three parties now measure this state, they get three bits with even parity *with certainty*. It is now clear how the strategies of the parties should be: If

you get 0 for input, you measure; if you get 1, you apply a *Hadamard* and measure. By the symmetry of the state, all that remains to be verified is that if *all three* parties apply a Hadamard, they always get odd parity. To verify this, let us first isolate the second and third Qbits from the first:

$$(\mathrm{H}\otimes\mathbf{1}\otimes\mathbf{1})|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle\left(\frac{|00\rangle-|11\rangle}{\sqrt{2}}\right)+|1\rangle\left(\frac{|01\rangle+|10\rangle}{\sqrt{2}}\right)\right) .$$

Observe now that

$$\begin{aligned}\mathrm{H}^{\otimes 2}(|00\rangle-|11\rangle) &= \frac{1}{2}\left((|0\rangle+|1\rangle)(|0\rangle+|1\rangle)\right.\\ &\qquad \left.-(|0\rangle-|1\rangle)(|0\rangle-|1\rangle)\right)\\ &= |01\rangle+|10\rangle .\end{aligned}$$

The two entangled states are swapped, and

$$\begin{aligned}\mathrm{H}^{\otimes 3}\,|\Psi\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle\frac{|01\rangle+|10\rangle}{\sqrt{2}}+|1\rangle\frac{|00\rangle-|11\rangle}{\sqrt{2}}\right)\\ &= \frac{1}{2}\big(|001\rangle+|010\rangle+|100\rangle-|111\rangle\big) .\end{aligned}$$

The parity is now always odd, as the game requests.

Chapter 7

The Needle in the Haystack: Grover's Algorithm

7.1 Motivation

This algorithm solves a very generic problem: Let f be a function from n bits to one bit, such that there exist only *few* (for instance, *one*) x_0 such that $f(x_0) = 1$; we call such an x_0 a solution, and it is the goal of the algorithm to find such a solution.

If the function is given as a classical black box, then the expected number of calls is of the order $\Theta(2^n)$. We will see that, if, on the other hand, it is a quantum circuit, roughly the square root thereof is enough: $\Theta(2^{n/2})$. Grover's algorithm [9] does not offer an exponential advantage, but it is very generic.

A typical application of Grover's algorithm is the exhaustive key search in a known plain-text attack to a block cipher: Here, the effective key length can be divided by 2. Generally, one-way function inverses can be computed with the same speed-up. Furthermore, collisions of generic hash functions, *e.g.*, from n to $n-1$ bits, can be found in time $\Theta(2^{n/3})$, which in this case is the classical running time, $\Theta(2^{n/2})$, to the power 2/3: The advantage is a little less dramatic since it is not a purely unstructured search problem. Generally, a naïve search algorithm for solving an *NP-complete* problem such as 3-SAT can be sped up accordingly, *i.e.*, from 2^n to $2^{n/2}$ steps, if n is the number of atoms involved. Note that this is still exponential, and many people believe quantum computers cannot solve NP-complete problems efficiently.

7.2 The Elements

The Grover circuit essentially consists of two unitaries, one of which implements f in the usual way (reversible gate plus phase kickback), whereas the other is independent of f. Let us start by defining this latter unitary. We first define a unitary Z through its spectral representation:

$$Z := |\mathbf{0}\rangle\langle\mathbf{0}| - (\mathbf{1} - |\mathbf{0}\rangle\langle\mathbf{0}|) \ .$$

Here, the term in brackets is the orthogonal projector to the orthogonal complement of the vector $|\mathbf{0}\rangle$. Geometrically, Z is a reflection with respect to $\mathrm{span}(|\mathbf{0}\rangle)$, and it can be written in short as

$$Z = 2|\mathbf{0}\rangle\langle\mathbf{0}| - \mathbf{1} \ .$$

Let now

$$A := \mathrm{H}^{\otimes n}\, Z\, \mathrm{H}^{\otimes n} = 2\,\mathrm{H}^{\otimes n}\,|\mathbf{0}\rangle\langle\mathbf{0}|\,\mathrm{H}^{\otimes n} - \mathbf{1} \ .$$

We define

$$|\Psi\rangle := \mathrm{H}^{\otimes n}\,|\mathbf{0}\rangle = \sum_{\mathbf{i}\in\{0,1\}^n} |\mathbf{i}\rangle \ ,$$

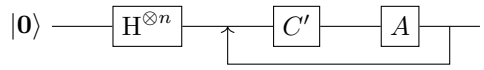
and

$$A = 2|\Psi\rangle\langle\Psi| - \mathbf{1} = |\Psi\rangle\langle\Psi| - (\mathbf{1} - |\Psi\rangle\langle\Psi|) \ .$$

The operation A is, therefore, a reflection at the line $\mathrm{span}(|\Psi\rangle)$.

The second unitary in the Grover circuit is C', mapping $|\mathbf{x}\rangle$ to $(-1)^{f(\mathbf{x})}|\mathbf{x}\rangle$: It is obtained from the usual reversible circuit via phase kickback.

7.3 The Grover Circuit



In order to figure out what this circuit does, we aim at linking the elements independently from f of the algorithm with f. More specifically, let us define the equal superposition of all solutions to f, and of all non-solutions. We can expect to be able to then write $|\Psi\rangle$, which is the equal superposition of *all* inputs, as a superposition of these two vectors: Let M be the number of solutions and $N - M$, where $N := 2^n$, the number of non-solutions of f. Then

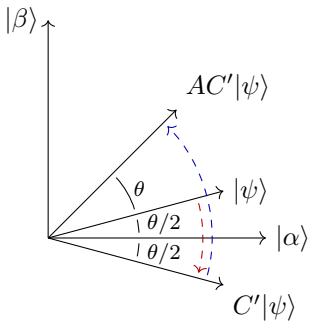
$$|\alpha\rangle := \frac{1}{\sqrt{N-M}} \sum_{f(\mathbf{x})=0} |\mathbf{x}\rangle \ ,$$

$$|\beta\rangle := \frac{1}{\sqrt{M}} \sum_{f(\mathbf{x})=1} |\mathbf{x}\rangle \ .$$

What makes the understanding and the analysis of the algorithm pretty simple is the fact that "all the action" going on in an a priori exponentially large space happens in the real plane spanned by α and β. Indeed,

$$|\Psi\rangle = \sqrt{\frac{N-M}{N}}\,|\alpha\rangle + \sqrt{\frac{M}{N}}\,|\beta\rangle\ .$$

In the interesting case, where M vanishes or is very small when compared to N, $|\Psi\rangle$ is close to $|\alpha\rangle$. Let the angle between the two be $\Theta/2$:



Both unitaries C' and A are now reflections at two vectors in that plane, C' at $|\alpha\rangle$, and A at $|\Psi\rangle$. The two taken together are a *rotation* in the plane by Θ. When this rotation is executed a suitable number of times, the resulting vector in the circuit is close to $|\beta\rangle$, and a full measurement yields a solution with high probability. Note that, when the rotation is continued, the probability shrinks again. Because of this property, Grover's algorithm is sometimes compared to a *soufflé*: If you do not take it out of the oven in time, it collapses again. However, the situation is not all that bad, since it then increases again. Indeed, the fact that all the action happens in the plane in question, and that within that plane, the state performs a uniform rotation, leads to a remarkable "emancipation" of the solutions in the face of the (much larger number of) non-solutions.

If the number M of solutions is known, the optimal number k of Grover iterations is roughly satisfies

$$k\Theta \approx \frac{\pi}{2}\ .$$

For small Θ, we have

$$\Theta \approx 2\sin(\Theta/2) = 2\sqrt{\frac{M}{N}}\ .$$

Altogether, a suitable number of iterations is

$$\frac{\pi}{4}\cdot\sqrt{\frac{N}{M}}\ .$$

The more solutions there are, the less iterations are required; that is quite natural. However, what if M is unknown? What if the algorithm is to test whether M vanishes or not (if, for instance, 3-SAT is to be solved)? In that case, the above "equilibrium" between solutions and non-solutions can be used: If you iterate, *i.e.*, rotate for long enough (the worst case is given by $M = 1$, so you vary from here), the probability of measuring a solution is roughly $1/2$. If in a sufficient number of repetitions, say 20 or 30, no solution is observed, then an event has occurred that would be extremely unlikely if there *are* solutions.

Chapter 8

Integer Factoring: Shor's Algorithm

The most impressive, aesthetic, and celebrated example of a quantum algorithm is Shor's algorithm for factoring integers and computing discrete logarithms in polynomial time [15]. These algorithms break most traditional public-key cryptosystems and have given rise to the field of *post-quantum cryptography*, *i.e.*, public-key schemes resisting quantum attacks. In the view of quantum cryptography discussed earlier, it is remarkable how colorful and rich the relationship between cryptography and quantum physics is.

8.1 Quantum Fourier Transform

We define a unitary transformation by defining its action on the classical basis of the n-Qbit space, and by giving an efficient circuit for it. The *quantum Fourier transform (QFT)* maps the state

$$|\mathbf{j}\rangle = |j_1 \cdots j_n\rangle$$

to

$$\frac{1}{2^{n/2}} \left(|0\rangle + e^{2\pi i \cdot 0.j_n}|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i \cdot 0.j_{n-1}j_n}|1\rangle\right) \otimes \cdots \otimes \left(|0\rangle + e^{2\pi i \cdot 0.j_1 j_2 \cdots j_n}|1\rangle\right) ,$$

where the expressions of the form $0.j_k j_{k+1} \cdots j_n$ are to be read as real numbers in binary expansion, where the j_k are its binary digits.

We give an efficient circuit for the QFT. Its elements are controlled-R_k transformation, where the R_k are generalizations of conditional phase flips.

$$\begin{bmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{bmatrix} .$$

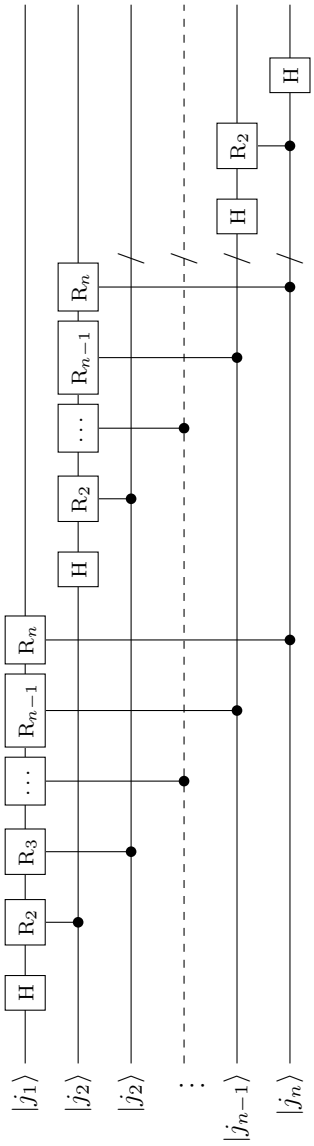


Figure 8.1: The circuit of the Quantum Fourier Transform.

The analysis of the circuit is quite simple since, for classical inputs, all controls are classical. So the wires can be computed through the circuit one by one. In particular, no entanglement is generated in that case.

Let us look at the first wire, with initial state $|j_1\rangle$. The *Hadamard* acts as a "self-controlled R_1" and generates the state

$$\mathrm{H}\,|j_1\rangle = \frac{|0\rangle + e^{2\pi i \cdot 0.j_1}|1\rangle}{\sqrt{2}} \ .$$

The next step is an R_2 controlled by the second Qbit, in state $|j_2\rangle$: If $j_2 = 0$, nothing happens, otherwise the phase of $|1\rangle$ is multiplied by the factor

$$e^{2\pi i/2^2} = e^{2\pi i \cdot 0.01} \ ;$$

all in all, this corresponds to a factor

$$e^{2\pi i \cdot 0.0 j_2} \ ,$$

and the resulting state is

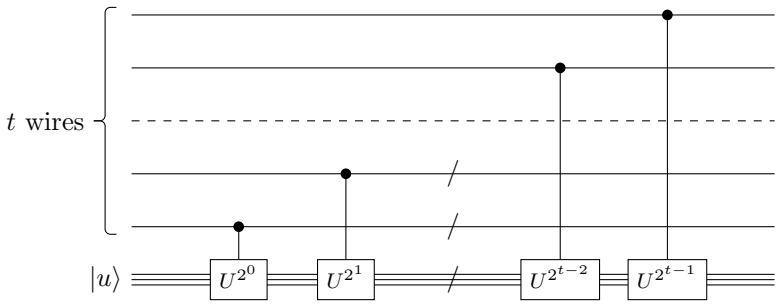
$$\frac{|0\rangle + e^{2\pi i \cdot 0.j_1 j_2}|1\rangle}{\sqrt{2}} \ .$$

This is the clue of the circuit: It addresses one by one the individual bit positions of the phase of $|1\rangle$ in order to yield the QFT of the input state $|\mathbf{j}\rangle$.

8.2 Phase Estimation

Consider the following problem: Let U be a unitary with eigenvector $|u\rangle$ and eigenvalue $e^{2\pi i\varphi}$. Assume that you are given $|u\rangle$ as well as the controlled-U^{2^j} for $j = 0, 1, 2, \ldots, t$. Find the first t binary digits of $\varphi \approx 0.\varphi_1\varphi_2\cdots\varphi_t$.

Let us assume the ideal case where $\varphi = 0.\varphi_1\varphi_2\cdots\varphi_t$, and consider the following circuit:

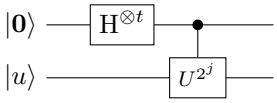


Let us consider the action of the controlled powers of U on the involved systems. The controlled-U^{2^j} does nothing if the control is $|0\rangle$, but if the control is $|1\rangle$, the action is executed:

$$\begin{aligned} |1\rangle \otimes |u\rangle &\mapsto |1\rangle \otimes U^{2^j}|u\rangle \\ &= |1\rangle \otimes e^{2\pi i \cdot 2^j \varphi}|u\rangle \\ &= e^{2\pi i \cdot 2^j}|1\rangle \otimes |u\rangle \ . \end{aligned}$$

Note that the last step is again the *phase-kickback* trick.

Let us now provide this circuit with a nonclassical input, namely, the usual $\mathrm{H}^{\otimes t}|\mathbf{0}\rangle$:



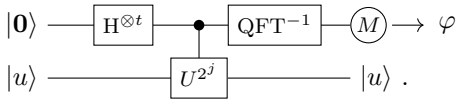
If phase kickback is used, then the state of the last group of wires remains unchanged and carries $|u\rangle$ throughout the circuit. The first wire ends up in state

$$\frac{|0\rangle + e^{2\pi i \cdot 2^{t-1}\varphi}|1\rangle}{\sqrt{2}} = \frac{|0\rangle + e^{2\pi i \cdot 0.\varphi_t}|1\rangle}{\sqrt{2}} \ ,$$

which is the last Qbit of the Fourier transform. It is, in fact, easy to check that the circuit produces the QFT of the string $\varphi_1\varphi_2\cdots\varphi_n$ is reverse order.

This observation solves our phase-estimation problem in the ideal case. In the non-ideal case, where φ does not terminate after t positions, the circuit can be set up for $t+s$ positions, and the first t positions are correct with probability at least $1-2^{-s}$.

In compact notation, we have the following circuit for phase estimation:



Here the notation means that if the first t wires carry the different (classical) bits of the binary expansion of an integer j, then U^j is carried out on the last wire: This is a compressed notation of what the circuit does. Note, however, that the state on the first t wires is *not* actually classical, as the notation insinuates.

8.3 Factoring

Let $N = pq$ be the product of two distinct odd primes. The order of $\mathbb{Z}_N^*$ is $(p-1)(q-1)$, and for many $x \in \mathbb{Z}_N^*$, we have

$$\mathrm{ord}(x) = \frac{(p-1)(q-1)}{2} \ .$$

Therefore, the ability to compute orders in $\mathbb{Z}_N^*$ implies the ability of factoring N.

Let $x \in \mathbb{Z}_N^*$, and let

$$U|y\rangle := |R_N(x \cdot y)\rangle \ :$$

U is the unitary corresponding to multiplication with x in $\mathbb{Z}_N^*$. The corresponding space is the state space of L Qbits, where $2^L \geq N$; for $y \geq N$, we define $U|y\rangle := |y\rangle$.

Note first that the controlled-U^{2^j} gate, acting as

$$U^{2^j}|y\rangle = |R_N(x^{2^j} y)\rangle \ ,$$

can be obtained by translating the classical "repeated squaring" method, in time

$$O((\log N)^3) \ ,$$

which is the asymptotic running time of Shor's algorithm. (This can be improved to essentially $O((\log N)^2)$ by using an asymptotically better algorithm for multiplication — based, ironically, also on the discrete Fourier transform.)

What are the eigenstates of U? Let, for $r := \mathrm{ord}_N(x)$ and $0 \leq s < r$,

$$|u_s\rangle := \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\frac{2\pi i s k}{r}} |R_N(x^k)\rangle \ .$$

Then

$$\begin{aligned} U|u_s\rangle &= \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\frac{2\pi i s k}{r}} |R_N(x^{k+1})\rangle \\ &= e^{\frac{2\pi i s}{r}} \cdot \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\frac{2\pi i s (k+1)}{r}} |R_N(x^{k+1})\rangle \\ &= e^{\frac{2\pi i s}{r}} |u_s\rangle \ , \end{aligned}$$

where we have made use in the last step of the r-periodicity of both functions involved. We conclude that phase estimation yields digits of the phase

$$\varphi = \frac{s}{r} \ ,$$

given that we know sufficiently many, *i.e.*, $O(\log N)$, digits, we can determine the period and, hence, the rational function, in particular r, which is the unknown.

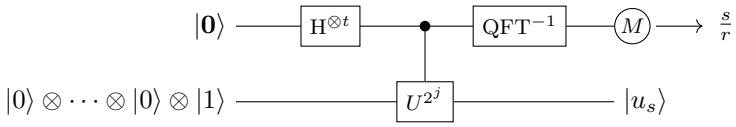
But how do we obtain one of the $|u_s\rangle$? In fact, we cannot: The definition of $|u_s\rangle$ depends on r; the argument is circular. But perhaps we do not have to — what is the equal superposition of all $|u_s\rangle$?

$$\begin{aligned}\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1}|u_s\rangle &= \frac{1}{r}\sum_{s=0}^{r-1}\sum_{k=0}^{r-1} e^{-\frac{2\pi i s k}{r}}|R_N(x^k)\rangle \\ &= \frac{1}{r}\sum_{k=0}^{r-1}|R_N(x^k)\rangle\sum_{s=0}^{r-1} e^{-\frac{2\pi i s k}{r}} \\ &= |000\cdots01\rangle\ .\end{aligned}$$

Therefore, we can simply take *this* state as the circuits input; measuring at the end leads us to a

$$\varphi = \frac{s}{r}$$

for *random* s, and in particular r.



Shor's algorithm is a very aesthetic and perhaps surprising connection between number theory and quantum physics — it has been called by *Klaus Hepp* the most exciting result in theoretical physics of the 1990s. It is also the main reason why nowadays adventurous sums are being invested in quantum technology. It is not obvious whether this is worth the effort, and whether that focus and resources could not be better invested to solve more pressing problems of mankind. Whereas it is a sportive challenge to build the first quantum computer, it binds substantial resources (as does football), and it may threaten our privacy. We refer the reader to the *epilogue* for more thoughts along such lines, and underline our admiration of Grover's and Shor's algorithms, because of their stunning beauty: We love them in the same way we love *Caravaggio*'s "San Giovanni Battista" or *Beethoven*'s Ninth Symphony.

Chapter 9

Epilogue: Information & Physics

The seminar "Information & Physics" was held parallel to the lecture. It has been broader than the lecture and covered a wide range of topics, including quantum cryptography, quantum foundations, interpretations of quantum theory, thermodynamics, its infamous "second law," the arrow of time, causality, and — sociological works concerning the role of science in society, and the role of society for science. It has been the goal of this — also very personal — closing lecture to make the connection between the technical (main) part of the seminar, and the sociological aspects; it was held on May 24, 2019 in ETH's ML E 12.

To wrap up, I (Stefan Wolf) have attempted to make a synthesis of the different parts of the event: How can we link and harmonize information and physics? How also to link the good old second law of thermodynamics, in the center of the first half of the semester, with these postmodern Bell correlations? How to link all this physics with the first week of the seminar, with Feyerabend's attacking myths on scientific practice yes, with Ludwik Fleck's studies on how scientific facts come into existence, I guess, but what about Hannah Arendt or Geoffroy de Lagasnerie? The latter task is eased by the day I am writing this, and its still warm impressions: First of May — labor day — first sunburn of the year!

The lecture was held on May 24, 2019. This write-up is under the impression of June 14 of the same year: "Wenn Frau will, steht alles still!" — the national women's strike: Second sunburn of the year.

The auditorium, one of the less romantic and more modern rooms in ETH's machine laboratory. If that place seems inaccurate: *Ernst Specker* held his seminar in one of the older wooden rooms there. It was held there the first

time I attended it (must have been in the early 1990s — the best years of my generation) — when it was held together with *Hans Läuchli*. It was also there, where my quantum information lecture was held a number of times, before it changed into an English event, doubled the participants, and moved to the sterile modern venue. Present in the auditorium was *Gilles Brassard*, the co-inventor of quantum cryptography and teleportation, what an honor. Present also was the "stoïc" philosopher and my friend and Gilles' student and stepson *Paul Raymond-Robichaud*, the only person with the courage to intervene the lecture with a valuable comment. By the way, Raymond-Robichaud is known for his and Brassard's *interpretation of quantum theory* called "parallel lives:" If a party making a measurement on a quantum system instantaneously splits into two (or more), and if the right split-off parties meet, then a local and realistic and deterministic view of Bell non-locality becomes possible. This idea was preceded by works of Deutsch and Hayden as well as jokes made on the topic by cryptographer *Matthias Fitzi*: "If the laboratory journals must be brought together in the end, what is non-local about the whole story?"

Present also was the rising "system philosopher," *Arne Hansen*, and the expert on *Michel Foucault*, this French provocateur through precise cold hurtful thought, *Ugo Balzaretti*. The lecture, for which I am about to take a deep breath to get it started, does not give him, Foucault, the honor to coin the whole lecture by giving him the first word — that first word *should* be given to someone else by any speaker, always, however —, but the softer, warmer, but no less sharp *Pierre Bourdieu*, about whom the speaker learned a lot from the sociologist and germanist *Benjamin Schlüer* (also present in the audience, present after the announced yet not implemented academic 15 minutes).

One of the main reasons for me to lecture at ETH, believe it or not (I do not know whether I do), has always been the joy of using a *black*board: The joy that seems to be forbidden at *Università della Svizzera italiana*. On the board, there are three words when the lecture starts: "ice" in blue, "opposing" in white, and "fire" in red.

The lecture starts.

"Just like the *state*, which takes away from the citizens their power to construct social reality, a *professor* has such a monopoly, limited to some hours a week, and some weeks per year." In the spirit of this quote by the French sociologist *Pierre Bourdieu*, I now seize half an hour or so — with your help: Actually, I need your help.

[One hears quite loud conversation from outside, an *apéro* is held, professor *Poulikakos* was seen there, so it was probably related to some machine-engineering activity. That is, the *apéros* at ETH are never directly related to *other* activities of the school, it is rather one of its major *stand-alone activities*, and they are really just drinking as drinking is. Maybe with a better

conscience, even better than when you just drink socially as opposed to home alone. After all, drinking with the explicit support of the *superego* must be barest of bad conscience. On the other hand, as the computational biologist and dearest of friends, *Manuel Gil*, once pointed out, you can spot the *real* drinkers at such a *apéro* by their drinking not at all or only like birds with very tiny mouths; the white wine is only discretely kissed as opposed to be made to really flow. On the same note, ETH is also famous for and proud of its (actually *her*, as you will see later) "*apéro* tourists," without whom the *apéros* would be way dryer even. After my own farewell lecture, there is no official *apéro*, after all, it is a *guerilla* farewell lecture in an occupied auditorium. The debriefing will be at *Safari bar*, as usual.]

Thank you all for being here, it is a great pleasure, and a great honor. We could all be also at the climate strike.

[Some attendees, for instance, *Andreas Wanner*, were actually at that manifestation before. Talking about him, a friend of mine a dearly missed at the lecture was the eminent mathematical physicist, *Jürg Fröhlich*, sharing with me the hometown of Schaffhausen — I had for a long time the suspicion that his generous and encouraging reception of my clumsy thoughts on physics were due to such local solidarity — or was it maybe my *not* being a physicist nor claiming to be. Do I have thoughts quite *ok* for an outsider. In any case, I would have loved my lecture to be honored by Jürg's murmuring and honoring critique; I am sure some of the things I says are "wirklich grosser Unsinn." My response: Of course, this is my language game here, my theatrical performance, bare of any claim of truth beyond that play. To entertain is the *only* goal.]

Thank you for being here. And I do need your help, because, as *Judith Butler* says, "even a monologue requires a structured space, a platform formed by people: if *one* person acts, *many* act." In this sense, thank you for acting with me here today, and also thank for having acted with me throughout the whole seminar. Sometimes, the audience was not that big, it was more a chamber play than a symphony. But it was always very nice.

[For that matter, also a *text* comes to life only through being read, being thought through, being opposed, being laughed at, being inspired or insulted by. Hence, also to you, dear reader, thank you for being here!]

My goal today is to wrap up the seminar now, to make a *tour d'horizon, une "vulgarisation non-vulgaire," comme le disait Georges Canguilhem.* The seminar had the title "Information" [written now to *blue* side of the board] and 'Physics" [*red* side]. Here, we have a first *opposition* — and "opposing" is the title of this lecture. I start by telling you the *history* of the seminar — that, as every "history," is just a collection of stories.

[This is in the "archaeological" style of *Michel Foucault.* What's better, even: what *other* way is there to speak about the essence of a story, an event, a tradition, than by going over its history? You may say: But has this history

not been twisted and turned? Maybe yes — but then, this is part of the phenomenon we look at. I would say that if you look at the present only, as it is very popular in science today, coming with a certain ridiculization of the old research through superficial admiration (gönnerhaft vom hohen Ross herab), then you are *really* exposed to self-glorifying manipulations by the protagonists acting today, waiting eagerly for their prizes and recognitions. I prefer the romanticized history of today's myths about scientific glory and supremacy.]

The story of the seminar cannot be told without telling the story of "ETH and me." [Sounds like: "I cannot tell Kurtz's story without telling my own." In *Apocalypse Now*, the continuation is then the remark that the story is "really a *confession*."] ETH was, when I joined *her*... I say "her": A student once told me that to him, ETH is like a jealous girlfriend, saying "No, you cannot get married to another girl while you would have exams with me instead," etc. For me, ETH was actually a *mother*. For sure it is a *woman*, there everybody agrees. One also says *"alma mater,"* right? It is definitely not only a school, for me, it was a *superego*. This here is a planet with only ETH on it, I call it "planETH."

[Here, I take into my hands a ball made out of burned clay, roughly the size of a handball, looking like the moon, rusty red, with one single building on in, the main building of ETH with a small lake Zurich next to it: The *blue* part of the planet. The piece of art, which is now floating in or behind Berne in the Aare river, was made by the artist *Heiko Schulze* from Kassel, whom Sonja and I met when we were there for the *documenta*. He presented such planets with single houses on them, and we asked him to produce "planETH." It will soon be replaced by a planet with a large beach chair and two drinks with straws, and a calm, sometimes wild ocean — not too *blue*, though.]

ETH was also, when I joined it, my clothing store. Today, when I arrived, I was in that clothing store that there is now in this building. At the time, in the late 1980s, there was not such a big choice in clothes, there was only one single type of t-shirts, maybe two colors or so [*blue* for sure; *red* there was not, or at least I did not buy it]. In my first two years in Zurich, this was where I bought my clothes. Today I was there and bought this [holding up a dark gray t-shirt with small text on it], it resembles most of those t-shirts I bought then, at the time. In a sense, the main change of ETH during that time, is reflected in the possibilities of what you can buy. Now, that store looks a bit like the fan store of *Bayern Munich*. What you can buy there today for me represents the biggest change ETH made since then: This marketing [I wanted to say: Merchandising], a very clear awareness of what image you want to project from yourself. Also the "media relations" were not so organized and streamlined at the time. Also the embodiment [that is: incorporation] of the critique, that you talk about of what you think of yourself is new in this form: ETH says here *"Loading Premium Education."* ETH was never exactly *modest*, but such

explicit self-praising is simply vulgar. [And, to some degree, exaggerated: The performance of participants of the same lectures and seminars I gave at USI and at ETH was well higher at USI, not because the students are brighter, but they have more maturity and free space and air to breathe. At ETH, my anarchic style brought people to follow other teachers' pressures and attend and study less. Using pressure to fulfill your goals, for instance, "premium education," is neither elegant nor noble, certainly not liberating. Was it not ETH's goal at its beginning to *free* people?

Another note on the aspect of incorporation of critique: On the occasion of the discussion of the film *The Matrix* on YouTube's *Filmkritik* channel, the smart young man with the fly quotes Slovenian philosopher *Slavoj Žižek*: In a sense, *"The Matrix"* can also be read as a *critique* of Hollywood itself — we, the consumers, are sucked out, they take our money and produce reality, dreams, movies. Incorporating also the critique of a system into the system simply makes it all more totalitarian. This text appearing in *vdf Hochschulverlag AG an der ETH Zürich* confirms the openness of the latter, and it underlines the readiness to incorporate even critique on incorporation of critique: *Gödel lässt grüssen. Und Hegel.*

The other major change ETH has gone through since I joined it is due to the introduction of the *Bologna* system that universities in all of Europe had to endure.

Back to me: I left my hometown *Schaffhausen* to join ETH because I wanted to learn something I could hold on to, I wanted to learn about *eternal truths independent of people.* I thought that this could give me security, that was my idea. In the context, I write here, on the *blue* side, the name of the pre-Socratic Greek philosopher *Parmenides.* He was the first to link the idea of *truth* with the idea of *eternity.* Before him, this link had not really been made, and truth was not seen as something that is eternal. But then, in Western philosophy, this connection was almost like carved into stone. [*Paul Feyerabend* has said that Parmenides misled Western thought for millennia — also an achievement, in the end.]

I wanted to learn about these truths, but I realized that all the interesting stuff going on at ETH was the talk about *people*, in my case mainly the talk about the mathematicians. There was, for example, this *Évariste Galois*, who got killed in a duel about a girl at the age of not even 30 years, and in the night before he died, he feverishly wrote down the theory of finite fields that no-one would understand for centuries. Then, there is the story of *Corneliu Constantinescu*, the tiny professor with the huge *tie* who would tell that story in his lectures, with the undertone *nota bene* of how harmful romance [that is: girls] were for mathematics, there were rumors about *Jürg Fröhlich*, this brilliant mathematical physicist, there were accounts of the overfull auditoria of *Paul Feyerabend*, rare in philosophy at ETH, otherwise more a *Feigenblatt*

[fig leaf] of the school. Unfortunately, I just missed his lectures at ETH, but I was there early enough to hear his stories.

And then, *Ernst Specker*, with whom I would like to continue now. Specker was a professor for mathematics, he was already retired when I joined ETH, but he was still offering, together with *Hans Läuchli*, the seminar. It was a *logic* seminar, and in a sense, I often saw it as a role model for this seminar here. [I add "Information," on the *blue* side of the blackboard, "/logic", also *blue*. In one of its last editions, the seminar was on *quantum logic*, and that was very inspiring for me.

Let us come to *this* seminar. The seminar has been on a whole number of topics, today we saw a certain overview of this. [There had been talks from different parts of the seminar, including on *Scott Aaronson*'s "NP-complete problems and physical reality," which explains the bottle of the soap *handy* on the table: The students had demonstrated an experiment using soap bubbles for solving a difficult computational problem, namely, finding "Steiner trees" using soap bubbles.] The content of the seminar is, actually, quite a mess. There used to be even cryptography in the seminar, classical cryptography, and then it goes on to other topics, ending up with sociology. This simply reflects my personal restlessness in terms of what topics I found interesting at different times. There was quantum cryptography, invented by the gentleman sitting there in the audience, *Gilles Brassard*. I am very proud you are here.

[One of my very early memories as a researcher was a lecture given by Gilles Brassard at ETH in 1994, when I was a Diploma (before *Bologna*) student investigating the security of the Diffie-Hellman cryptosystem, given that the underlying problem, the "discrete logarithm," is really hard. Gilles has a very long beard, he was drinking liters of water, and he was explaining the use of quantum mechanics for information processing. He has been introduced by Ueli Maurer as a *Mozartian* "Wunderkind," having skipped every second grade at school. Gilles seemed like from another planet, and he never completely lost that touch for me, to this day, when he looks so much more earthed.]

There was quantum information, then quantum foundations; the seminar became more and more foundational. Causality — we talked about causality, space-time causality —, and then there was this first week of the seminar on *sociology*. What is that? How can that even be linked to the rest? It is one of the goals of this concluding lecture of mine to link this sociology part to the rest of the seminar. I do this by talking about *oppositions*, as insinuated already by the title. I synthesize some oppositions, I resolve some of them, find a way to get rid of them, I take a clear side in some of the oppositions, and I make some of the oppositions even stronger instead of resolving them. In fact, oppositions can be very fruitful and even generate energy. Some authors, like *Byung-Chul Han*, professor for philosophy in Berlin, describe our times, and the difficulty of living in our times, as a *lack of opposition*: We live in times of pure positivity. There is no otherness; it has simply been sent away [or, as

discussed above: erased through incorporation]; it is not accepted anymore. Reality then becomes *totalitarian*.

[The noises from the "otherness" around the auditorium, the *apéro*, become louder: The effect of the white wine is audible.]

I have a quote on that, from the *Wochenzeitung*, by Raoul Zelik: "In the capitalist desert of the real, there is no exterior anymore. The source of the sadness of our times is that 'there is no alternative.' " On the other hand, he then goes on, capitalism needs *expansion* to survive. In the end, what you get in consequence is that what is produced: all this "new" stuff, is always the same, it is a copy of a copy of a copy. It get rarer and rarer that you do something really *different* from what there already is.

Let us now conquer this sadness and allow for opposition, even *celebrate* opposition, and, in particular, allow for *otherness*. The otherness of *Parmenides of Elea* [on the *blue* side] that avoids his eternal-truth obsession become too totalitarian, is a contemporary of his, that is *Heraclitus of Ephesus*. [His name is written to the *red* side.] He was another pre-Socratic philosopher, just like Parmenides. The characterization of the opposition between the two through what I had written on the top of the board, *ice* versus *fire*, is by Nietzsche in his text "Die Philosphie im tragischen Zeitalter der Griechen." He wrote that Parmenides and his thinking style is like *ice*: An ice-cold logician, everything is based on logic. *Heraclitus*, on the other hand, is the *fiery* physicist, fire! Physics is the basis of everything, change is the basis of everything. Probably you know a quote he is famous for: "You cannot step into the same river twice."

Let us try to get some neutral territory here (if there is any neutrality): *Jeanne Hersch* [her name is written in white between the two pre-Socratics] has been a professor of philosophy at *Université de Genève*, and she has said that the opposition between Parmenides and Heraclitus *coined the entire history of philosophy*. All philosophers can somehow be put into the *blue* or the *red* region. And then, I thought, well if this is true for the history of philosophy, then it might be no less true for the history of physics. In fact, many debates have taken place in physics that somehow fit into this picture. We start with a debate between *Newton* [his name is written in *red* below that of Heraclitus] on the one hand and *Leibniz* on the other. The debate was about space-time: For Newton, space-time is what you *start* from, it is pre-given, like a stage which is also already there even if no play is happening on that stage. In this sense, physics is here really the basis of everything that happens, all the logic, everything. Leibniz did not agree with that: For him, space-time was first of all only *relational*, and neither absolute nor fundamental: If there is nothing happening, there is also no space-time: The theater stage emerges only when there is a play, along with that play.

In the face of this "early modern" chapter of the opposition in consideration, it is fair to say that the course of physics thereafter was almost entirely

in favor on Newton. That includes also *Einstein*, who modified space-time compared to Newton, but it is still fundamental, even if nothing happens [no masses are around], there is still [the flat] space-time [of special relativity]. On the other [*blue*] side, there were some physicists belonging there, most notably *Ernst Mach*: For instance, "Mach's principle" states that inertial forces are *relational*, that is, they do not come from acceleration towards absolute space, but with respect to other objects, such as the fixed stars. As we will see below, this view has become today more and more popular.

The *postmodern* chapter of the opposition is expressed — and we have talked a lot about that in the seminar — on the *blue* side, *John Archibald Wheeler*'s "It from Bit." This means that physical reality emerges somehow from information, from pieces of information, that is the basis. The opposite, on the *red* side, we have also discussed this a lot, the opposite position, uttered by *Rolf Landauer*, "Information is physical:" Here, Landauer sees physics as the basis, and information can always only be seen in the context of its physical realization, and physical laws thus have consequences for information processing. And *the* physical law that is normally named here, in the context of Landauer's principle, is *the second law of thermodynamics*.

The second law of thermodynamics is very dark. The *Woody Allen* movie "Husbands and Wives" paraphrases the law: The protagonist Sally talks to her analyst and says she did not know why her marriage broke apart, and then she says: "Yes, actually I *do* know why, it is because of the second law of thermodynamics, 'everything turns into shit sooner or later,' this is my phrasing, not the *Encyclopedia Britannica*," she says.

[Loud laughter, for the first time *inside* the auditorium without white wine!]

Question by Paul Raymond-Robichaud: "*Donald Knuth* proposed that computer science could perhaps be seen as a pure science, which does not depend or is linked in any way with reality." [This is an interesting thought. Note, however, that "doing computer science" is certainly linked with reality, as a social enterprise, see below.] Alright, let us put him here into the neutral zone; like in ice hockey, we have a neutral zone here, and Knuth perhaps belongs there. ["Knuth" is written in white between Wheeler and Landauer.]

Back to the second law: As dark as it may be, it is important for us to orient ourselves in time: It gives time flow a direction, which other physical laws do not. Probably, there is also a *use* of the second law for the creation of memory. Our memory makes time very asymmetric for us, we remember yesterday, but we do not remember tomorrow, so the second law must be somehow involved here. This is maybe a bit like *friction*: It is seen as a problem, slowing down all movements and so on — but, actually, without friction, we could not move at all. We use friction when we walk, or when we ride a bike. Maybe [the entropy increase of] the second law has, for our memory, a similar role and is not that pessimistic, after all.

Maybe we could also say that because *the* example for Landauer's slogan "Information is Physical" is always the second law that *the second law is actually not physical*, that it in the end belongs to the *blue* side.

Another phenomenon that maybe also belongs to the *blue* side are *quantum correlations*, about which we heard a lot in the seminar. These are pieces of information (the weirdness becomes manifest when looking at the information aspect of the phenomenon), and these pieces of information are so strangely correlated that it means, in the end, that the same piece of information pops up at different locations at the same time — it is extremely weird. If we try to do a synthesis from this (these gentlemen there, Gilles and Paul, also made a synthesis): Maybe the phenomenon questions Newton's assumption that space-time is fundamental: Quantum correlations can be seen as an argument for seeing space-time causality as not that fundamental. Now, if space-time is not fundamental, like temperature, then it has to emerge somehow. "Emerging" means that is appears only in the *macroscopic* world, and it is not there in the microscopic world. Does it perhaps, like temperature, somehow emerge on the macroscopic level? [Links are drawn on the board, underlining the idea that thermodynamics might hold the key to understanding or resolving the measurement problem — which, however, might also be more of a *linguistic* than a physical problem, as suggested by Arne.]

I want to show you another synthesis: On the *blue* side here, we have these quantum correlations, completely "ignoring" physical distance, and on the *red* side the unavoidable decay of everything: I found the following in a book, "Stil und Moral," by *Lukas Bärfuss*, an author from Thun but living in Zurich. He wrote: *"We dominate space, but time dominates us."* [The first part is written on the *blue* side, the second the the *red*.] I cannot think of a more beautiful summary in a single sentence of the whole seminar, or at least of the tension between these two types of results here. Bärfuss is a storyteller, a fiction-writer — let us follow that line a bit, let us try to find some *stories* behind all this research.

We start here, on the *red* side. The second law of thermodynamics was discovered by *Sadi Carnot*, whose father was an officer of the French army, which was one of the organizations having early specimen of a pretty new invention, *James Watt*'s steam engine. These engines could be studied to figure out the law, that is, Carnot's version of the law characterizing the efficiency of such engines in function of the temperatures of the heat baths involved. [*Jürg Fröhlich* has underlined that Carnot has an H-index of *one* since this was his only publication; he was not even 30 years old.] Then there is the version of *Clausius*; you know that the Clausius Street is not very far from here [and the first dozen of times I stayed in Vienna, that was at *Boltzmanngasse* — you see, the law follows me; or do *I* follow it? [Little anecdote: Everyone using taxis in Vienna to go to Hotel Boltzmann or the Physics Institute knows that Boltzmanngasse is, like Boltzmann himself was, "bipolar," split in two.

The reason is the American embassy in front of which the street is completely blocked for cars, by "security" reasons, of course, the magic word terminating any reasoning and deliberation. At the same time, security would not even be hard to achieve, as *Noam Chomsky* put it: There is a simple way to avoid terrorism: *Do not engage in it.*] Back to Clausius: "Heat does not flow spontaneously from cold to hot." In the next sentence, he says, roughly: "Therefore, we all die, since in the end, the universe will be in a state where no temperature differences exist anymore." *Kelvin*'s version of the second law reads: "From one heat bath alone or from radiation alone you cannot get useful energy" — "...except if you are are a plant," then, you can do photosynthesis: also mysterious. For me, the central figure linked to the second law is *Ludwig Boltzmann*. He developed a modern view of the second law, realizing it is much more general than only connected to steam engines. He related it to probabilities and counting; he was an atomist, he believed literally in *atoms* that could be counted, arranged, etc., unlike *Mach*. The story goes in Vienna that Boltzmann had been mobbed by Mach, and that Boltzmann's suicide had been also related to that mobbing by Mach.

Boltzmann came from a very Catholic family, had a unruly life, was bipolar, then his tragic suicide in Duino in Italy, where he was on holidays with his family, just days before the lectures would start. There is a strange link (not only in Boltzmann's case), a strange link between the lives of the people involved and the pessimistic statement of the second law.

To the other, *blue*, side: Let us contrast this with the discoveries here, and about the discoverer of these mysterious correlations quantum theory gives rise to, *John Bell*. Bell worked as a particle physicist at CERN, Northern Irish, so very Protestant, and what he is famous for now was his *hobby*, he even worked on weekends: "I am a quantum engineer, but on Sundays I have principles," that is on Sundays he was working on the foundations of quantum theory. So hard actually that in the end his health was affected by all this hard work.

The contrast between the results we oppose here could not be stronger. You do not need to trust experimentalists to believe in the second law of thermodynamics, its signs are *everywhere*. Whereas, on the other hand, quantum correlations are very subtle, not easy to reconstruct in the lab. But then, at the same time, they are *optimistic* results, they have so many applications, you can even found a company based on it, as it was done in Geneva, you can do cryptography using the correlations: promises are attached to it, hopes, and certainly a big fascination. Among other things, it allows for *quantum teleportation*, also co-invented by *Gilles Brassard*.

[Idea for a novel: The young, attractive, very successful *Joanne Bellman* works in the field of quantum correlations and their uses for cryptography. She works for a Geneva-based enterprise, has a tanned physicist boyfriend, working at CERN, with a muscular body, not overly smart and with unlimited praise and admiration for Joanne. She has just been promoted to vice-president of

the company. At the party for celebrating that, she meets in a dark corner of the world a much older man, dark but sharply handsome, *Ludwik Bolts*. He has studied physics ages ago, suffered a psychosis following a long obsession with the second law of thermodynamics. In the end, the darkness of perspectives has been extremely liberating to him, he found lightness in his not having an influence beyond the *here and now*. Joanne falls in love with Bolts in a way she has never experienced before, the intensity overwhelms her, she loses the ground below her feet, her boyfriend, etc. She even starts to develop conspiracy theories about the correlations that are the basis of her company's business to be based on manipulated experiments. Also, she questions fundamentally the most profitable part of her company, selling devices, for thousands of Swiss francs, producing just "randomness." What a first-world, unnecessary, luxury product is that? Moreover, in endless nightly discussions with Bolts, the two have established the view that decay processes are, actually, not related to loss of information, but to the excessive accumulation of it. So in the end, devices producing randomness accelerate decay. This is, Joanne thinks, a symbolic reflection of the perverse style having been established on the planet, sometimes referred to as "globalized capitalism": Those who rule live in luxury while the others must deal with the randomness they produce: The first world needs many *Maxwellian demons*, an average first-world citizen holds, metaphorically, about 60 slaves, as computed by a German economics Professorin *Evi Hartmann*. In euphoric disillusion, Joanne quits her job, throws her smartphone into the Rhône, ends up only reading books that are at least as old as Bolts, no news consumption anymore, no consumption at all, starts working in the fields, enjoys the fresh air and the sunshine and realizes she is happy for the first time in her life. Somehow, she feels true and real and fulfilled. Bellman and Bolts spend happy years in harmony, and after his calm peaceful death, she throws his ashes into wild winds and waters, smiling and crying and singing at the same time.]

In summary, there are very different groups of results. Do they represent a difference between the inner workings of the people involved? Can we maybe say (and this is now really speculative [and in "a fiction-writer thinking style *à la Bärfuss*"]). I talked before about *the Freudian superego*, that the *blue* group of results represent the "superego of Physics," whereas the *red* side, a bit darker, represents the "id of Physics." Or maybe this is more accurate, maybe the *blue* side represents the "Eros of Physics," whereas the *red* is the "death drive of Physics" — a very strong Freudian tension. What is the meaning of this opposition (maybe also culturally)? Let us go back to *Bourdieu*, whom I quoted right at the beginning of the lecture: "The opposition between *Nature* (*red*) and *Culture* (*blue*), between negligence on the one hand and well-behavedness on the other, is the Freudian opposition between the *id* and the *superego*."

Let us look at the context: Where did this come from, from what kind of society? On the *blue* side, this is quite clear: Bell's results come from the second half of the 20th century, a time of complete *"boom"* [the word is written to the *blue* side]. On the other, *red* side, Boltzmann's Vienna and Europe, was in total *decline* [word written to the *red* side]. Let me quote here what *Erich Fried* said about Freud: He said he did not want to diminish Freud's insights, but discovering *repression (Verdrängung)* in Austria is not very hard. In the same sense, I ask the question: "How Viennese is the second law of thermodynamics?" ["Vienna" is written to the *red* side]. On the other side, what do we have? Bell is working at CERN, and CERN is really the dream machine of physics, the *Hollywood of Physics*, in Geneva [Geneva is written to the *blue* side].

I ask the question: In which sense do scientific results also express *views*, views of society and views of individuals? In the first week of the seminar, we heard a lot about that. For instance, we talked about *Ludwik Fleck*: For him, the background of results are "thought collectives," *C. G. Jung's* "archetypes" are a similar notion, *Žižek* talks about ideology, and *Foucault* used the term "episteme": Episteme is the totality of what a society knows and believes, and out of which all new science grows.

Allow me now to apply this view to a topic we discussed in the seminar, the "interpretations" of quantum theory. Allow me to wrap up the quantum measurement problem in a nutshell: Say you have polaroid sunglasses with the property that if a light ray hits it, then exactly half the light goes through. Then, the question is what happens if a *single* light particle hits the sunglasses, a "light atom," a photon, that cannot be split into two. The possibilities then are that either it does *not* go through *at all*, or it goes through *completely*. But it definitely cannot be divided in half. This is a little weird, and the question is how, where, and by whom it is decided whether the particle goes through or not. Here, different "interpretations of quantum theory" say different things. If Jürg Fröhlich were here today, he would probably become angry now, and a little loud, he would perhaps say: "Quantum theory does not have to be *interpreted*, it has to be *understood*!" Still, there are different readings of the theory, such as the *Copenhagen interpretation* of quantum theory, the traditionally "standard" one going back to the founding fathers of quantum theory. ["Copenhagen" is written to the *blue* side of the board.] According to that view, outcomes of measurements in quantum theory (that is, whether the photon goes through or not) are *random*. This can be related to societies where "free will" is highly regarded, in Calvinist Geneva for example. *Max Weber* noted that the spirit of capitalism comes out of Protestant ethics. Again, note that the *blue* side is the Protestant side, where the *red* is the Catholic one. *Zwinglian* Zurich is also one of the places where there is still a strong belief in this view. On the other side, for example, there are *deterministic* readings of quantum theory such as "Bohmian mechanics" ["Bohm" is written to the *red*

side]. Bohm was actually a very convinced Marxist, and this Marxist background was perhaps one of the reasons why he preferred a determinist reading of the theory. By the way, Marx was not such a strong determinist, whereas *Engels* was. Finally, we have other, more "postmodern" interpretations of quantum theory: take *Everett*'s, more populistically phrased, "many worlds" or "parallel lives." It means that if the light particle hits the sunglasses, then there are two worlds, one in which it goes through, and one in which it does not. I put it here to the neutral zone on the boards, since it is, like Bohm [on the *red* side], *deterministic*; it has, however, also this "neoliberal" component, this multioption component. [Paul is shaking.]

Let me explain, I would like to do it by first giving you a parallel: The "Luhmannian" sociologist *Elena Esposito* pointed out that *probability theory* was actually invented at the same time as the modern *fiction novel*. The common background in society was what Luhmann called "Realitätsverdoppelung," doubling reality.

[Probably, society as a whole overstepped here a limit that every child crosses, at the age of about 4 years. *Ernst Specker* once described the experiment where *Kasperle* hides his chocolate in some place. Then a thief comes, takes it from there and puts it somewhere else. Then Kasperl comes back and says that he would like to put his chocolate to a safer place. Where will he go and get it? If you pose this to very young children, they will say that he goes to where the chocolate now actually *is*. When they are older, they say he goes to where he *(erroneously) believes it so be*, where he had put it. This is the age at which children also start lying. Does this mean that when society invented probability and novels, it also invented *modern politics*?]

The "reality doubling" gave people the possibility not to see a novelist as a lier, but as someone creating a piece of art. At exactly the same time, Pascal invented probability calculus. (Another such coïncidence may have been the connection between the invention of *book printing* and of *national states*.) So we may ask: The appearance of the many-words interpretation, is it perhaps not related to the emergence of the Internet, offering people the possibility to live in parallel realities and societies that co-exist at the same time but do not communicate at all? Which, by the way, is also true to some degree for the different "interpreters" of quantum theory. They also have their camps between which there not too much communication is going on; they all have their reality, their truths.

We have arrived at the notion of *truth*. ["Truth" is written in the neutral zone.] What makes us believe what is a truth, and what do we relate it to? On the blue side, we have Parmenides and the idea that truth is eternal. And also objective: The truth is true for everyone. The idea is popular, also here at ETH, that what physical theories actually do is to yield better and better approximations to truth, making truth better and better accessible.

There is also another view, namely, that truth is more *cultural* and more *subjective*, more *constructed*. One of the first to question the Parmenidean view of truth was *Friedrich Nietzsche*: He wrote a text entitled "On truth and lies in an nonmoral sense." He writes: "So what is truth? A mobile armada of metaphors. Truth are illusions of which one has forgotten that they are illusions." And then he elaborates. What I observed is often brought forward as supporting the truth of a theory as a proper, precise mathematical structure, and so on. But *why* is this an argument for truth? [The thought that nature is simple is popular and more or less formally put as *Occam's razor*, *Solomonoff induction*, universal probability, etc. It cannot, however, be justified really. One gets stuck in some kind of *Humean problem* here.] Nietzsche: "All the wonder about the natural law that we so admire lies in the mathematical precision of time and space representations. Those, however, we produce in us, with the same necessity with which the spider cocoons." So views of space and time for Nietzsche do not come from outside. Nietzsche breaks here with the idea of eternal truths, of truth that is independent of people. Then, the question is: How come, quite some time after Nietzsche, today is the idea of eternal truth still so persistent? *Paul Feyerabend*, who was a professor here at ETH, wrote that Parmenides has in fact misled Western thought for millennia by his linking of truth and eternity. How is that possible? For *Bourdieu*, this is actually a sociological question: *How can a timely activity produce timeless truth?* For Bourdieu, it means that society has installed mechanisms that guarantee eternity or at least the long duration of these truths. There are several ideas linked to this, and the first idea is that scientific production is a somehow *superior* human activity. [This idea is normally not made explicit but presumed unquestioned.] In his inaugural lecture in 1980, Paul Feyerabend criticized this by saying: "Man entscheidet sich für oder gegen die Wissenschaften so, wie man sich für oder gegen *punk rock* entscheidet," uttered with his nice Austrian accent.

Let us return to Bourdieu's thought: If you want to guarantee that your timely activity produces timeless truth, then you need *procedures* to guarantee that. One of the procedures installed is *"peer review,"* which is definitely today the standard procedure in science. Peer review often supports and promotes what looks similar to what is already around. Obviously, peer review has certain effects on where science goes; it sometimes acts as a kind of *censorship*, since you are perhaps less free to write what you want to write than to say what you want to say. We have also seen in the seminar that peer review can lead to *sexism* and to *nepotism* in research; it leads to *conservation*, and it has a *normalizing* effect, it introduces *norms*. A norm always means there is an "inside" and there is an "outside": Either you satisfy the norm, and you are inside the circle, or you are outside the circle. *Bourdieu* writes that the education system is like a *Maxwell demon*: The official narrative is that it gives everyone the same chances and so on, but he says (and he is French, and maybe

the French system here is somewhat more accentuated in this sense) that the opposite is actually the case, that a selection is made about who fits into the system of the *écoles normales* [and this selection us more based on *habitus* than ability: the most important parameter may be whether your father is a *normalien*, etc.], and who should be excluded. In any case, the "insiders" then run the entire country: "The ruling ideas are the ideas of the ruling class." This quote is not by Bourdieu, it is by Marx and Engels, from their *Communist Manifesto*. Just one page later we read: "Communism abolishes eternal truth, religion, moral, instead of repairing them." This remarkably complements what we have said about truth. So let us politicize the opposition a bit: Communism (*red* side of the board) versus Neoliberalism (*blue* side). And maybe anarchism would be more here, in the neutral zone ["A" in a circle is put in the middle in white].

In the end, we have now left the field of science in a stricter sense and arrived at society as a whole. *Michel Foucault*, in his inaugural lecture at the *Collège de France*, asked what the concept of truth meant in a society. Where does it come from? What is its role? He says discourse in a society has to be regulated: Discourse is something very dangerous for those in power, and it must be closely monitored, controlled, regulated, and normalized by certain mechanisms. He says that there are essentially three mechanisms to doing this. The first one is *taboo*: Don't say that! You do *not* say that! The second one is the *sanity/madness separation*: Yes, you can say that if you want. But if you do, we will put you in some kind of institution. [Gilles laughs.] The third and most important one is the *true/false* separation: This is a truth obsession — which can also be a *truth oppression*. This mechanism is particularly efficient because it leads to *self-censorship*: This is exactly the effect of the superego, this is what the superego does; it is implanted into everyone, repeats "do not say wrong things, just say true things," and you have people controlling themselves. This reminds a bit of *Friedrich Dürrenmatt* saying that Switzerland is a prison, and everyone is a prisoner and a guard at the same time. For Foucault, conceptions of truth go hand in hand with power structures. This also means, put a bit more positively, that truth can also be *oppositional*. Foucault wrote: "To write is to fight and to oppose and to map." He saw himself also as a cartographer. [I am not sure that was true; it was for *Juliamia Stirnemann*.]

What field of science is oppositional? Is sociology oppositional? I quote here the French sociologist *Geoffroy de Lagasnerie*. He denies that there are oppositional and non-oppositional fields. Rather, there are only *oppositional versus conservative ways of behaving within a field*. It is not important what field you belong to, but whether you promote an emancipatory project in that field. There is an interview of his online in which he starts by saying that there actually only thinkers of the left. Of course, on the right side there are also people saying words with their mouths, "mais ce ne sont pas des penseurs, ils

sont juste traversés par des pulsions ou des idéologies. Les penseurs sont les gens qui questionnent fondamentalement se qui se passe, et ils sont toujours vu de faire partie de l'extrème gauche, comme Bourdieu, Foucault, Sartre, Camus, etc." Lagasnerie says that it is actually not an obligation to fight what you do not agree with, after all, we are born into a society that we do not choose. But he says that if you enter the field in what he calls *"symbolic production,"* if you become an artist, or a writer, or a scientist, then there is not really anymore a neutral position for you. Then, you are expected to ask yourself how society is going, and how you with your actions contribute to that. Science today, however, does actually quite the opposite: Science today likes to promote of itself an image of *neutrality*. He says this neutrality is like the neutrality of Switzerland in WWII. I come from Schaffhausen, I told you, there is actually a story about that from Schaffhausen. There was the company *SIG* exporting a lot of weapons to the Germans. People said that in Schaffhausen, people were working for the Germans six days a week, and on the seventh day, they went to church to pray for the victory of the allies: That is also "neutrality." The other story I heard just two days ago is this: First, Schaffhausen was bombed by the Americans in 1944, on April 1, there were significant casualties. Officially, it was a mistake. [After all, Schaffhausen is located north of the Rhine river.] Besides the weapons, Schaffhausen's IVF was also exporting tons of cotton to the Germans that was in some way being used for the bombs, for the triggers. Then the Allies said "you cannot export that much cotton to the Germans, as a neutral country." Ok, was the reply, "we'll just export shirts from now on." So they then exported shirts that of course were made out of cotton and could easily be used for the bombs again. The allies again: "Come on, you cannot export tons of shirts to the Germans, this is not neutral." Then they said: "Ok, we'll limit it to 100'000 shirts a year" — and it seems the shirts now were 20 meters big: "Let us bomb them!"

Just as science is never neutral, also *technology* is never neutral. Let us take an example: *artificial intelligence*. I start with a quote by *John von Neumann*. The context is that they were developing the nuclear bomb using the very early computers. He said: "We are creating a monster to change history, if there is history left." It was usually believed that the whole phrase concerns the bomb, but actually only the second half of the sentence did — "if there is history left." The first —"We are creating a monster to change history" — concerned the *computers*. He felt maybe how invasive the technology could become to our daily lives. I would like to connect here with what Foucault said in his lecture on neoliberalism in his lecture on "bio-politics." This also is very invasive. The intellectual basis of neoliberalism originated in *Freiburg im Breisgau* in Germany in the 1930s, under the name of *ordoliberalism*. Again, we have the idea of creating a *norm*. In this case, the norm is the market logic, which invades all aspects of our lives. In *classical liberalism*, the ideal is a disappearing state, whereas in *neoliberalism*, the state comes back with full

strength, not as an economic player, but to guarantee the imposition of the market logic to regulate all aspects of society.

The market logic then becomes totalitarian, and is no longer equal to other logics. Today, train lines are closed because they are not profitable. But profitability was hardly a driver when the railways were built, rather the dream to connect people. The same is true for social insurances, AHV and IV. How absurd is the idea to cut spending to the maximum in a rich country like Switzerland.

[Probably, the underlying psychological mechanism is that we know our welfare is based on theft or, thermodynamically speaking, our "orderly" society creates disorder, and death, in so many places of the world. The bad conscience triggers the superego, which suggests that we should at least not "waste" the money. So, in the end, we treat the underprivileged of our own societies bad as well, and feel some justice. This goes hand in hand with the feeling of "self-righteousness" (Selbstgerechtigkeit), a very common feeling in Switzerland: If we are better off, then with good reason. I am sorry about the others, but their deplorable state is because of their own behavior, be it that they "live in a country with poor working morals," be it that they live here but do not work or, even worse, smoke pot. In the latter case, "invalidity insurance" refuses to pay your rent or job intervention, because you refuse to collaborate, as *Sonja Ramseyer* told me.]

The totalitarian information technology combined with the omnipresence of neoliberalism leads to an explosive mix. For instance, surveillance: That is omnipresent and can become more and more omnipresent. The cryptographer *Yvo Desmedt*, when he visited us in Lugano recently, said that actually, it is a *scandal*: Secret services are supposed to do surveillance of the governments in the name of the people, not *vice versa*. The role of the secret service is to check which government member frequents which lobbyist, and so on.

What we see sociologically in the face of the surveillance technology is a dissolution of *private space*. The separation between *private* and *public* space was important for *Hannah Arendt*, whom we heard of also in the seminar. We oppose *private* [written to the *red*] to *public* [*blue*]. Again, the private space obviously dissolves, there is not even private space in *Ibiza* anymore. For Arendt, the public space is the space where you *act*, where you do politics. These were Roman ideas she took over, that were certainly accurate 2000 years ago, but today, this has completely changed. A first change is that if private space dissolves, then everything becomes public and, with that, political. Ironically, this is accompanied by the emergence of new political subjects that just do not accept the *necessity to go to public space to be political*, like *Chelsea Manning* or *Anonymous*. They seize the right to make politics without being a public figure. [One of the threats for the political "apparatus" today is the exaggerated public interest for the individuals involved.]

Other novel political subjects also *politicize* state membership, they do not accept a priori their being born into a state and having to accept its rules without ever being asked about this. They also evade prosecution, which sharply separates their actions, for instance, *Assange*'s or *Snowden*'s, from civil disobedience *à la Thoreau*. Anonymity and escape have become political instruments. At the basis is, again, *not accepting* to be forced into a state contract you did not choose in the first place, but you were forced into. Let us hear again here a literary comment on it, this time by an Austrian, *Thomas Bernhard*: "Der Staat hat mich, wie alle anderen auch, gefügig gemacht und aus mir einen Staatsmenschen gemacht, einen reglementierten, registrierten, trainierten, absolvierten, pervertierten und deprimierten, wie die anderen." In short, the state swallows you, and you have no choice.

For *Geoffroy de Lagasnerie*, the "nation-state" exists between bookprint (as discussed, book print enabled the creation of myths broad enough to give birth to social constructs such as nation-states) and the *Internet*: It is so much able to create new spaces such that the state could look like a very old idea soon. A first symptom of this would be these new subjects we talked about that fundamentally criticize state membership and the obligation to be loyal to a state into which you are randomly thrown. Lagasnerie: "To have as a mental background the whole world, and to free oneself from forced-upon memberships, this could be the axes of the art of revolt arising today and into which those participate that manage to define themselves as *citizens of the world*." The latter is how normally *Anonymous* addresses the public. That phrase is the end of Lagasnerie's book "Snowden, Assange, Manning — The Art of Revolt."

I am almost at the end of this lecture, I would like to come back to myself a bit: What should we do? What can I do? My script here, you cannot see that, has become quite *red* in the meantime, *blue* has virtually disappeared. Maybe I can draw now a separation between *me* [*on the red side*] and *ETH* [*blue*].

[Liberated and liberating laughter.]

According to *Hannah Arendt*, speaking in the public space is also *action*. But maybe action should not be limited to speech. Let me put here to the *red* side "action," and to the *blue*: "thought," Parmenidean logic.

You, Andreas, were at the climate strike today. On those topics, politicians like the German liberal Lindner like to hear themselves say in calm [you guessed it: *blue*] voice, in the face of student protesters' hot [*RED!*] anger: "Things are very complicated, they do not understand the economic connections, etc." This is a blunt lie. Things are *not* complicated at all. Also the fact that lives are being lost in the Mediterranean See: This is not a complicated situation. It is just *hard* to do something about it because it has a price. Wittgenstein: "Das Einfachste ist das Allerschwierigste, weil nicht eine Schwierigkeit des Intellekts, sondern des Willens überwunden werden muss." The *simplest* is

the *hardest*, because it is not a question of "intellect" [written to the *blue* side], but a question of "will" [*red*]. With "will" is not meant here the neoliberal "free will," but the Kantian will, the will to do the right thing, not the one for choosing some product.

You realize, I advocate for the *red* side. But what is bad with thought? After all, we studied quantum correlations, we do not build weapons. Ok, the correlations can be used in cryptography, but still, we do not essentially build weapons. So what can be wrong about it? What can be so bad about fundamental research? In order to say something about that, let me end by quoting again *Lukas Bärfuss* [who is, as we now know, the recipient of the 2019 *Georg Büchner-Preis*].

"Wie nicht wenige unter Ihnen war auch ich bisher der Ansicht, die Lektüre eines kulturkritischen Essays..." — *in our context maybe: 'The study of Bell inequalities,' for instance...* — "...sei dem Weltfrieden zumindest nicht abträglich" — *'does not threaten world peace:' obviously, studying Bell inequalities does not endanger world peace* — "aber ich habe die Seiten gewechselt" — *I have changed sides; why?* — "Falls Sie nicht einsehen, welche moralische Sauerei Ihre Lektüre darstellt," — *he means here the lecture of his, that very text; transposed: If you do not see how scandalous your sitting here now, and listening to me is, and my speaking of course as well...* — "stellen Sie sich bitte folgende Situation vor: Eine gutgenährte, wohlhabende Person, Ihnen gar nicht unähnlich, verschlägt es in ein sagen wir afrikanisches Flüchtlingslager, in dem gerade Cholera ausgebrochen ist. Menschen schreien, sterben. Doch statt zu helfen, sucht sich unser fiktives *ich* eine einigermassen ruhige Ecke und beginnt sich an der Lektüre von Rilkes 'Sonnetten' zu ergötzen." — *again, taken to our context: You are at a refugee camp in Yemen, pure horror, you are there, but instead of helping, you sit into a corner of the camp and start to read* Deutsch*'s 'Fabric of Reality.'* — "Sie müssen zugeben, dass dieses Verhalten zumindest moralisch fragwürdig ist, und sie müssen auch zugeben, dass wir im Grunde alle in einer etwas ruhigen Ecke eine Flüchtlingslagers leben. Die Entfernung macht das Elend perspektivisch kleiner, und nur Idioten glauben, das sich entfernende Auto werde tatsächlich zum Punkt." — *distance does not make suffering smaller* — "Sie sehen, die Lektüre literaturkritischer Essays ist in diesen Zeiten moralisch nicht zu rechtfertigen, und deshalb gehe ich mit gutem Beispiel voran und höre hier nun auf." — *you see, he says, the study of, allow me to transpose, the study of Bell correlations in times coined so sharply by action of the second law of thermodynamics is morally not tenable, so I want to be a good example and stop right here."*

Chapter 10

Bibliography

[1] John S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(3), 1964.

[2] Charles Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, page 175, 1984.

[3] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993.

[4] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997.

[5] David Deutsch and Richard Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 439(1907):553–558, 1992.

[6] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, May 1935.

[7] Avshalom C. Elitzur and Lev Vaidman. Quantum mechanical interaction-free measurements. *Foundations of Physics*, 23(7):987–997, Jul 1993.

[8] Walther Gerlach and Otto Stern. Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld. *Zeitschrift für Physik*, 9(1):349–352, Dec 1922.

[9] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pages 212–219. ACM, 1996.

[10] Grete Hermann. *Die naturphilosophischen Grundlagen der Quantenmechanik*. Abhandlungen der Fries'schen Schule. Verlag "Öffentliches Leben", 1935.

[11] Alexander S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

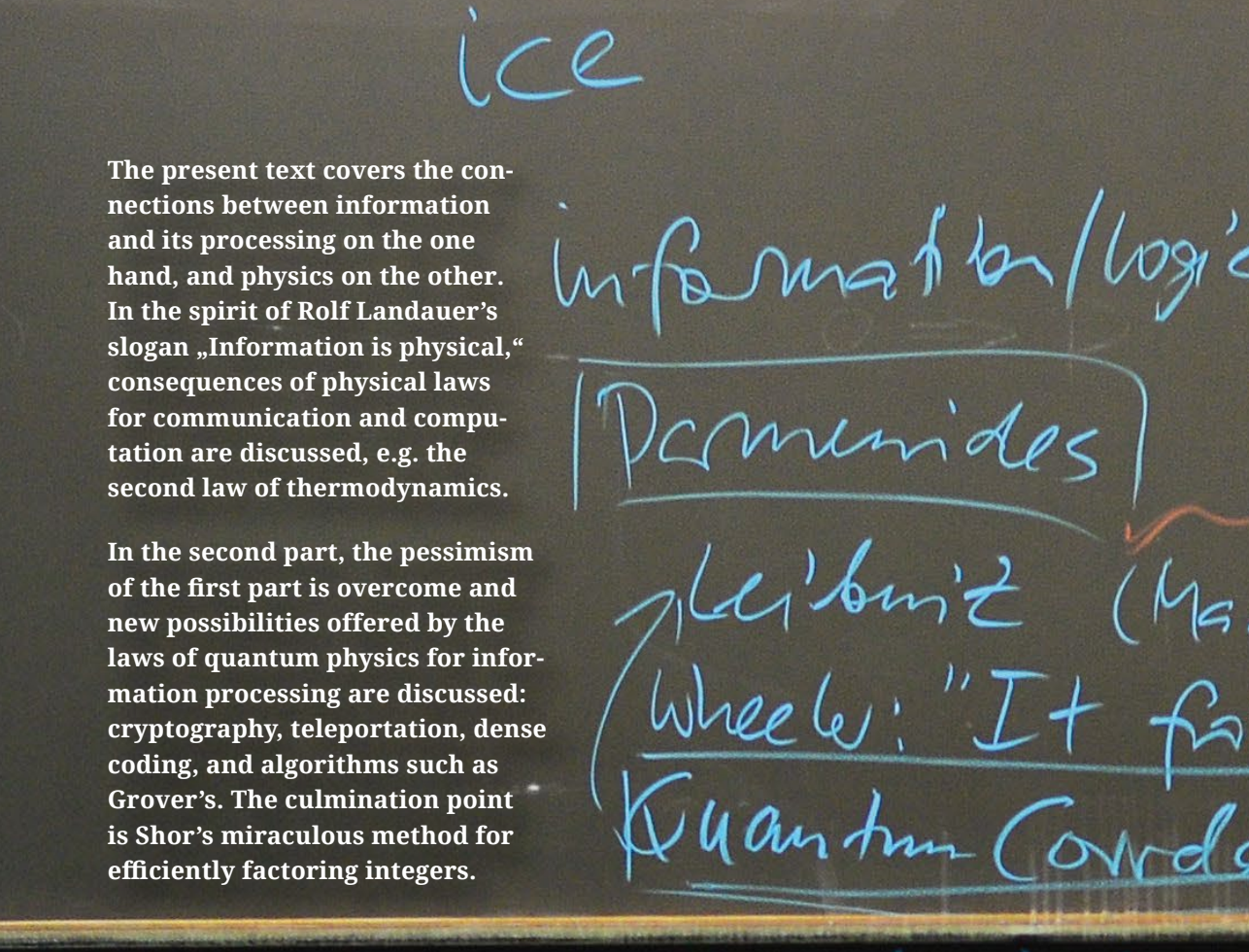
[12] N. David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65:1838–1840, Oct 1990.

[13] N. David Mermin. Copenhagen computation: How I learned to stop worrying and love Bohr. *IBM Journal of Research and Development*, 48(1):53–61, 2004.

[14] Karl Popper. *Logik der Forschung*. Mohr Siebeck, Tübingen, 11 edition, 1934.

[15] Peter Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Review*, 41(2):303–332, 1999.

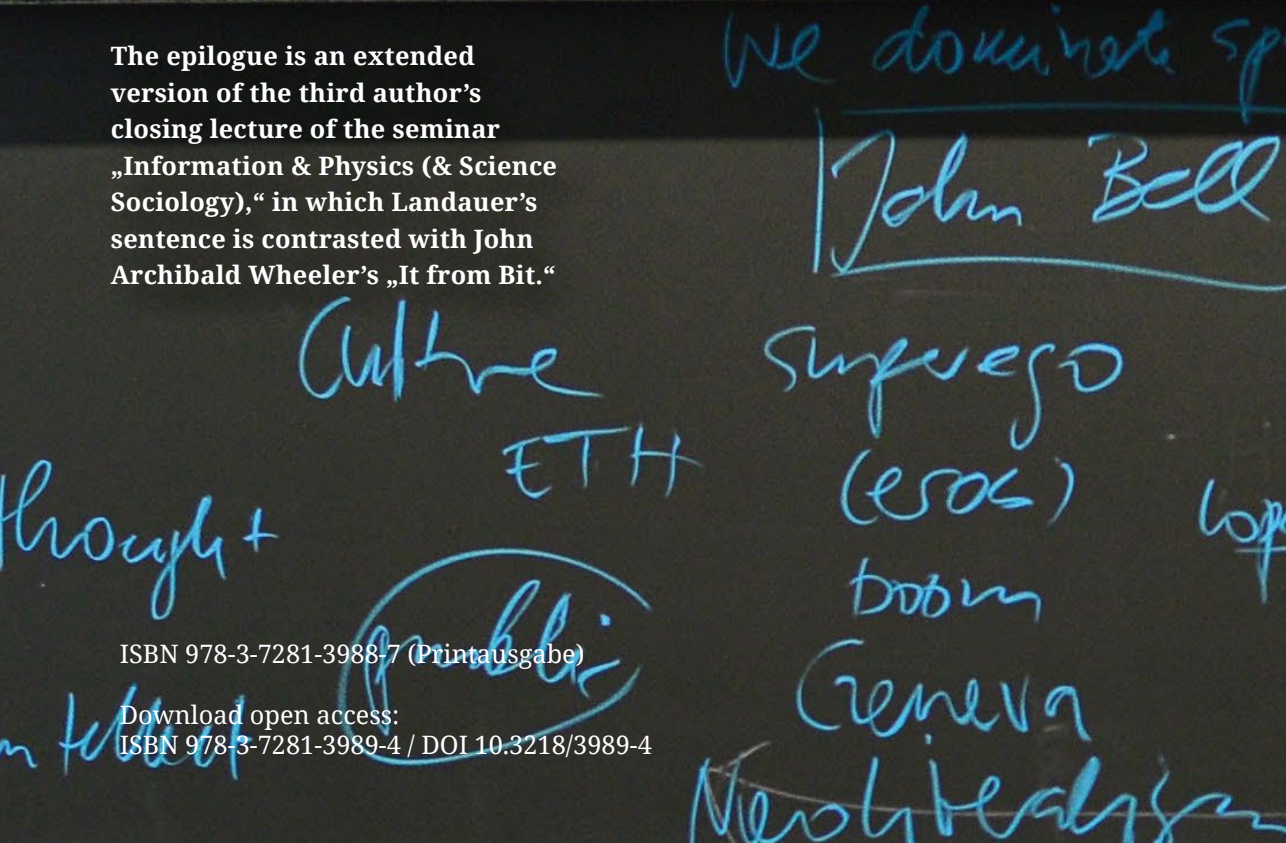
[16] Otto Stern. Ein Weg zur experimentellen Prüfung der Richtungsquantelung im Magnetfeld. *Zeitschrift für Physik*, 7(1):249–253, Dec 1921.



The present text covers the connections between information and its processing on the one hand, and physics on the other. In the spirit of Rolf Landauer's slogan „Information is physical," consequences of physical laws for communication and computation are discussed, e.g. the second law of thermodynamics.

In the second part, the pessimism of the first part is overcome and new possibilities offered by the laws of quantum physics for information processing are discussed: cryptography, teleportation, dense coding, and algorithms such as Grover's. The culmination point is Shor's miraculous method for efficiently factoring integers.

The epilogue is an extended version of the third author's closing lecture of the seminar „Information & Physics (& Science Sociology)," in which Landauer's sentence is contrasted with John Archibald Wheeler's „It from Bit."



ISBN 978-3-7281-3988-7 (Printausgabe)

Download open access:
ISBN 978-3-7281-3989-4 / DOI 10.3218/3989-4