Technische Universiteit Delft
Faculteit Elektrotechniek, Wiskunde en Informatica
Delft Institute of Applied Mathematics

Quantumlogica en de EPR-paradox

(Engelse titel: Quantum logic and the EPR paradox)

Verslag ten behoeve van het
Delft Institute of Applied Mathematics
als onderdeel ter verkrijging

van de graad van

BACHELOR OF SCIENCE
in de
TECHNISCHE WISKUNDE

door

LAURA MOLENAAR

Delft, Nederland
augustus 2014

BSc verslag TECHNISCHE WISKUNDE

"Quantumlogica en de EPR-paradox"
(Engelse titel: "Quantum logic and the EPR paradox")

LAURA MOLENAAR

Technische Universiteit Delft

Begeleiders

Dr. E. Coplakova

Dr. K.P. Hart

Overige commissieleden

Dr. D.C. Gijswijt

Prof.dr. J.M.A.M. van Neerven

Drs. E.M. van Elderen

Augustus, 2014

Delft

Contents

1 Introduction: Is Logic Empirical?

> *"The first process, therefore, in the effectual study of the sciences, must be one of simplification and reduction of the results of previous investigations to a form in which the mind can grasp them."*
>
> J. C. Maxwell

In the early 19th century, the famous mathematician Karl Friedrich Gauss discovered - parallel to other mathematicians - something remarkable: the fundamental axioms of geometry as defined by the Greek mathematician Euclid in his book 'The Elements' did not suffice to describe all objects. Euclid had tried to build geometry from as few assumptions as possible, almost 2100 years before Gauss. Until then, these axioms were regarded as being true, but most mathematicians had some issues with the fifth postulate, the *parallel postulate*:

> If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles. *Euclid [7]*

Although these five postulates formed a complete set of axioms ('complete' in the logical sense), mathematicians had trouble with this last postulate. The reason for that was that it was much more complex than the other four postulates. Mathematicians tried for hundreds of years to derive this postulate from the first four, unsuccesfully. But Gauss thought the postulate to be problematic for another reason: the parallel postulate was insufficient to describe the geometry of spheres, hyperboles and ellipses. *Empirically* gathered information indicated that spheres needed a fundamentally different kind of geometry than flat surfaces. Along the Euclidean geometry, the field of non-Euclidean geometry as discovered by Gauss was explored. In other words: after empirical evidence, it was decided to alter Euclidean geometry for curved surfaces like the sphere.

In the case of logic a similar event occurred: practice radically changed theory. Logic, contrary to geometry, was not at all seen as an empirical science. For a long time, logic was based on a few axioms that were taken as undoubtedly true. These axioms were regarded as purely mathematical and had little to do, supposedly, with how nature works. However, when the field of quantum mechanics emerged, scientists saw that there were some strange concepts that did not obey the rules of the classical logic that was conventional at the time. In particular, the classical law of distributivity,

$$p \wedge (q \vee r) = (p \wedge q) \vee (p \wedge r),$$

turned out not to be true for quantum mechanical systems. This was a shocking discovery in two ways: Firstly, classical logic had to be altered for quantum mechanical systems, and secondly, axioms in logic could be proven to be insufficient for some systems in an empirical way, exactly like what happened in the field of geometry. As it happened, nature disproved distributivity for quantum mechanics[1]. We will see an example of why this distributive law failed in Chapter 2.1.

These two examples show that mathematics can be viewed as an empirical science. Physical phenomena can lead to drastic changes in theory. However, this interaction between practice and theory can lead to paradoxical situations. In quantum mechanics, lots of paradoxes were discovered (and refuted) by scientists, since quantum mechanical objects behave in a fundamentally different way from the objects we encounter in daily life. One such paradox is the famous

[1] For a more detailed account of this analogy, see Putnam [17]

EPR-paradox, named after physicists Einstein, Podolsky and Rosen.
This thesis is about the non-distributive quantum logic, and how it deals with this EPR-paradox in particular. First, we give a short overview of quantum mechanics in Chapter 2 and in Chapter 3 we will give an introduction to the mathematics of quantum mechanics. In Chapter 4, we will explore Hilbert spaces, as they play a crucial role in the mathematics of quantum mechanics, and in Chapter 5 we will connect these Hilbert spaces with the lattice of propositions from Chapter 3. After that, we discuss superposition and entanglement, which both play a vital role in the EPR paradox, in Chapter 6. In Chapter 7, we give a formal construction of quantum logic. Finally, in Chapter 8, we will examine the EPR paradox on a mathematical level.

2 A Short Introduction to Quantum Mechanics

Quantum mechanics describes the properties of (sub)atomic particles. The theory behind these particles is quite complex. Not only mathematically, but also conceptually. If we want to have a proper view of the mathematics of quantum mechanics, we need to introduce some of these concepts. This will be done in this chapter.

2.1 Measuring Quantum Mechanical Systems

Quantum mechanical systems behave in a fundamentally different way from classical (physical) systems. In classical systems, it is (at least in theory) possible to know everything there is to know about a system. If one has, for example, a celestial body, one is in principle able to measure its exact position, speed, temperature et cetera. In quantum mechanics, even if one knows everything the theory can tell about, for example, a particle, one still cannot predict the outcome of some experiments. This is remarkable: It is impossible to predict certain aspects of a system. As it turns out, it is only possible to give *statistical information* about the possible results. We will see later on why the statistical view is the best we can do.

The problematic measurement of quantum mechanical systems is one of the most intrigueing examples of the difference between classical and quantum mechanics. Suppose we measure a particle at some position x. Where was the particle before we measured it? Quantum mechanics is not able to predict where such a particle would be if we had not measured it. As we have said earlier, it will only give us a statistical view in the form of a probability distribution of all possible positions. Over time, several answers have been proposed to the question of pre-measurement position. One could say: the particle was at x. But then why couldn't we predict that the particle was at x? Our theory would then be incomplete; there are certain aspects of reality that our theory fails to predict. This is the *realist* view of quantum mechanics.
Another approach is to say that the particle was not at a fixed point. The act of measurement forced the particle to choose one of its possible positions, and it ended up at x. The *orthodox view*, is now widely accepted among physicists.
The uncertainty in measurements turns out to be characteristic of quantum mechanical systems. It expresses itself for example in Heisenberg's uncertainty principle:

$$\Delta x \Delta p \geq \frac{\hbar}{2}.$$

Where $\hbar$ is the reduced Planck constant, approximately $6.63 \cdot 10^{-34}$ Js. In words: One cannot know the precise location *and* momentum of a particle simultaneously. The more precise the position (Δx) becomes, the less precise the momentum (Δp) can be known.

Heisenberg's uncertainty principle, and therefore the statistical nature of the quantum mechanical quantities, is a consequence of the *wave-particle duality*. This principle is an important aspect of quantum mechanics (and particle physics). Any particle must be viewed as both a particle and a wave. Just as it is impossible to precisely locate a wave on a piece of string, it is impossible to assign a position to a sub-atomic particle. In the same way, it is hard to assign a wavelength to a wave on a string[2] when you give the string just one sudden jerk, since the wave is not periodic. This principle can be translated to the waveforms of particles such as the photon. The more precisely you try to determine its wavelength (momentum), the more spread out the position becomes.
The fact that measurement is problematic has as a consequence that the distributive law fails for quantum mechanical systems. Let us take for example, a particle in a box, that can move

[2]The wavelength is directly associated with the momentum via the De Broglie formula, $p = \frac{2\pi\hbar}{\lambda}$, where $\hbar$ is again the reduced Planck constant.

on the real line. Let us define the following propositions:

p : 'The momentum of the particle lies in the interval $\left[0, \frac{1}{6}\right]$ ',
q : 'The position of the particle lies in the interval $[-1, 1]$',
r : 'The position of the particle lies in the interval $[1, 3]$',

where all numbers have been normalized (divided by $\hbar$). Then

$$p \wedge (q \vee r) = (p \wedge q) \vee (p \wedge r)$$

is not true. We know that $p \wedge (q \vee r)$ is true, but $(p \wedge q) \vee (p \wedge r)$ is not: Because of Heisenberg's principle, $\Delta p \Delta x$ must be greater than or equal to $\frac{1}{2}$. But this is not true for the expression $p \wedge q$. If it were true, p and q would both have to be true themselves. But then we bump into Heisenberg's uncertainty principle: We have a system where $\Delta p = \frac{1}{6}$ and $\Delta x = 2$, so that $\Delta p \Delta x = \frac{1}{6} \cdot 2 = \frac{1}{3} < \frac{1}{2}$. So the right hand side is not true. However, the left hand side is true. because there we have $\Delta p \Delta x = \frac{1}{6} \cdot 4 = \frac{2}{3} \geq \frac{1}{2}$. So indeed, the distributive law is not universally true for quantum mechanical systems.

2.2 Superposition

Quantum superposition is a fundamental principle of quantum mechanics. It expresses the convexity of the set of states of a quantum mechanical system. But before we can explain what this means exactly, we have to know what we mean when we speak of 'states'. A detailed (mathematical) description is given in Section 3.1, but here we will assume that a state is a vector in the (vector) state space consisting of the positions, momenta, spins etc. of all the particles in the system. That is, it describes the status of the system. We denote such a vector as $|v\rangle$. This is the bra-ket notation as used by Dirac, see [5]. A state that can be represented as a mix of two states, i.e.,

$$\alpha|u\rangle + \beta|v\rangle, \qquad \alpha, \beta \in \mathbb{C}, \text{ with } |\alpha|^2 + |\beta|^2 = 1,$$

is a superposition of these *pure* states $|u\rangle$ and $|v\rangle$. With the coefficients α and β the probabilities of being in one of the corresponding states can be calculated. The probability that the system is in state $|u\rangle$ is $|\alpha|^2$ and the probability that the system is in the second state, $|v\rangle$, is given by $|\beta|^2$. This makes it clear why we have put the second constraint on the coefficients.
As it turns out, every quantum mechanical state can be expressed as such a mixture of states (where α or β may be zero) and so the set of states is convex.

If we have an orthonormal basis of states for our state space, say, $\{\phi_i\}$, then every pure state Φ is expressible as

$$\Phi = \sum_i k_i |\phi_i\rangle, \quad \text{where } k_i \in \mathbb{C} \text{ and } \sum_i |k_i|^2 = 1.$$

Note that this is something *different* than the mixed state. This state is *pure*. Every pure state can be expressed as a combination of basis vectors, and every mixed state is a linear combination of these pure states.

An example: A system consists of a single electron, that has either spin up ($|\uparrow\rangle$) or spin down ($|\downarrow\rangle$). Spin is an intrinsic property of particles. It is often described as angular momentum, but really it has no 'classical' counterpart to which it can be compared. Spin has a direction (here: up, down) and it is *quantized*; it can be only of a size that is a multiple of $\hbar/2$ (but we will work with normalized vectors to make the notation friendlier). The set of basis vectors in our example is $\{|\uparrow\rangle, |\downarrow\rangle\}$. The electron is always in a superposition of these states.
Suppose we have the following description of its pure state:

$$\frac{1}{\sqrt{2}}|\uparrow\rangle + \frac{1}{\sqrt{2}}|\downarrow\rangle.$$

So the electron has spin up with probability 0.5, and spin down with probability 0.5. Before measurement, however, the electron is in both states at once. We cannot tell in which state the electron is before we measure it, and when we do, it 'loses' the superposition and collapses into one of the two states, each with probability 0.5. This problematic measurement plays a large role in the EPR paradox as well, see also Section 2.5 and Chapter 8.

2.3 Entanglement

Another phenomenon that occurs frequently in quantum mechanics is the entanglement of multiple particles. If two (or more) particles are entangled, their states cannot be expressed independently. To be more precise: The probabilities corresponding to the possible states of a particle that is entangled are no longer independent of the probabilities of states of the other particles it is entangled with: One state describes both particles. The occurence of this phenomenon entails that measurements performed on a single entangled particle have effect on the states of all particles it is entangled with.
Suppose we have prepared a system with two entangled particles in such a way that after measurement one particle has spin up and one particle has spin down, but we do not know which particle has spin up or down. Recall that *before* measurement, both particles are in a superposition of spin up and spin down. If we measure one of the particles and it turns out, for example, that the first particle has spin up, the other particle *instantly*[3] gets spin down. Note here that the distance between the particles has no effect on the entanglement itself. If we send one particle to the other side of the universe (assuming that the universe has 'sides') the result would not be any different.

This might sound paradoxical: How could the particle 'know' the state of the other particle? And what's more, how can it adopt a state without being measured itself? It is difficult to understand this intuitively; even quantum physicists Einstein and Schrödinger did not believe this at first. However, entanglement has been shown to be a real physical feature of particles. A more mathematical account of entanglement will be given in Section 6.2. Like superposition, entanglement plays a large role in the EPR paradox as well. This will be clarified in Chapters 2.5 and 8.

2.4 The Principle of Locality

The principle of locality states that a physical system can only be influenced by its direct surroundings. In systems we encounter in daily life, this seems very reasonable to assume. For example, objects in each other's proximity experience gravitational force, and this force subsides as the distance between the objects grows.
However, in quantum mechanics, locality is sometimes violated. For example, particles in entangled systems can be separated by huge distances and still influence each other. We say such systems are *nonlocal*. Entanglement is a nonlocal phenomenon, and thus independent of distance. Nonlocality plays a major role in the refutation of the theory of hidden variables, which was proposed by Einstein, among others, as a solution to the alleged incompleteness of the theory of quantum physics. Here, by incompleteness we mean incompleteness in the non-logical sense of the word: there are elements of reality for which the theory does not account. One argument for this incompleteness of quantum mechanics was given by Einstein, Podolsky and Rosen in 1935. They formulated a paradox, now known as the EPR paradox, which we will

[3]This really means instantly. There is no time difference between the first particle being measured and the second particle adopting the opposite spin. Although this seems to violate the principle that nothing can travel faster than the speed of light, it does not. Note that before measurement, we do not know anything about the spin of either particle, and we cannot influence it either. This means that no information travels faster than light, since no information is exchanged.

discuss in the next section and in Chapter 8.

2.5 The EPR Paradox

When the field of quantum mechanics was still young, it was strongly criticized. This is no surprise, as quantum mechanics entails a lot of strange phenomena and concepts. One of those critics was physicist Albert Einstein. In 1935, together with his colleagues Boris Podolsky and Nathan Rosen, Einstein wrote an article on an alleged paradoxicality in the theory of quantum mechanics. They thought the description of reality as given by quantum mechanics could not be considered complete. For them, the condition for completeness was that

> 'every element of physical reality should have a counterpart in physical theory.'

That is, to everything we 'see' should be an explanation from the theory. This condition would be satisfied when the following criterion holds:

> 'Without in any way disturbing a system, we can predict with certainty the value of a physical quantity.'

(both from Einstein, Podolsky and Rosen [6])

In quantum mechanics, it is assumed that the description of states by wave functions is a complete description of the physical reality of the system. EPR tried to show that this assumption together with the above mentioned criterion leads to a paradox:
Suppose we have two physical systems, I and II. Let $\{|\varphi_+\rangle, |\varphi_-\rangle\}$ be a complete, orthonormal set of vectors for the state space of system I and $\{|\psi_+\rangle, |\psi_-\rangle\}$ a complete, orthonormal set of vectors for the state space of system II. This can be read as two particles in two boxes (I and II) that can have either spin up (+) or spin down (-), for example an electron and a positron. The sets of vectors corresponding to each system describe these systems completely. Assume that we managed to entangle the systems[4], such that after some time, the state of the joint system is given by

$$\Psi = \frac{1}{\sqrt{2}} \left[(|\varphi_+\rangle \otimes |\psi_-\rangle) + (|\varphi_-\rangle \otimes |\psi_+\rangle) \right]. \tag{1}$$

This is mathematical notation for a system where there is a 50% chance that the particle of system I has spin up and the particle of system II has spin down, and a 50% chance that it is the other way around. Quantum mechanics cannot predict which state you will get when performing a measurement, but it tells you how the spins of these particles are correlated. Now, the probability that both particles are in the same state after measurement is 0, so if we have measured one particle (for example in the spin-up state), we immediately know the state of the other particle (namely, that it has spin down). We may now separate the particles by a large distance, and therefore we can by measuring the state in system I determine the state of system II *without* perturbing the second system. This means, according to Einstein, Podolsky and Rosen, that the states $|\psi_\pm\rangle$ have an element of reality (see the criterion above). Einstein, Podolsky and Rosen thought this action-at-a-distance to be untenable. The orthodox position says that the measurement results would be created by the experimenter, which is problematic. They refuted the orthodox position and adopted the realist view: one particle really had spin up and the other particle had spin down, we just did not know about it.
The value of the state in system II is not known before any measurement on system I is completed, but this value must have had the same element of reality before the measurement has

[4] An example of creating two entangled particles is the decay of a neutral pi meson (π^0) into an electron and a positron (e^- and e^+). The pi meson has spin zero, so the electron and positron are in the singlet configuration as in equation 1.

taken place, because a measurement on system I cannot produce any physical effect on system II (locality), and thus it cannot change the reality of a physical quantity in that system. So the element of reality that system II had before the measurement must be the same it has after measuring. Every time we measure the particle in system I to be in the up state, we know immediately that the particle in system II is in the down state, and vice versa. From this, we deduce that the compound system actually is in a mixture of two different states: The states $\Psi_1 = |\varphi_+\rangle \otimes |\psi_-\rangle$ and $\Psi_2 = |\varphi_-\rangle \otimes |\psi_+\rangle$, both with probability 0.5. Note that this is *different* from the state as described in Equation 1, since the state of Equation 1 is pure. This is a contradiction.

In other words: Because we can determine the characteristics of system II without directly measuring it (that is, by measuring system I) it must contain an element of reality. Therefore, quantum mechanics ought to describe those characteristics. But quantum mechanics does not determine the state the particle in system II is in without measuring it. The compound system consists of two entangled states and therefore, we cannot describe it completely. Thus, quantum mechanics must be incomplete. We will see in Chapter 8 how this paradox can be resolved.

3 The Mathematics of Quantum Mechanics

Before we can go into any detail, we must construct a mathematical account of what it means to perform an experiment. How can one represent a (quantum) physical system and the physical laws that interact with the system mathematically?

3.1 Propositions, States and Observables

We represent a *yes-no experiment*, an experimental question that can be answered by 'yes' or 'no', by a pair (A, Δ), where A is an *observable* and Δ is a real interval. An observable is a physical quantity that can be possessed by a certain system. *Any* experiment can be reduced to a number of yes-no experiments. One could examine, for example, whether observable A has a value in the interval Δ: Is the position (an observable) of a certain particle in the range $[0, 2]$ (an interval), given that it moves on the real line? An experiment can confirm this question, by taking the interval and checking whether the particle is present or not. Another example is 'Does the spin of this electron point in the direction of the positive y-axis?'. If the question deals with a smaller interval or multiple intervals, several yes-no experiments can be performed sequentially[5]. From such yes-no experiments, more complex experiments can be built. Intuitively, the *propositions* of a physical system can be viewed as (an ensemble of) yes-no experiments.

Before we come back to the propositions and examine the structure of the set of these propositions, we will define a few concepts.

We will call a set M *partially ordered* if there exists a relation $\preccurlyeq$ satisfying

1. $\forall\, a \in M : a \preccurlyeq a$
2. $\forall\, a, b, c \in M :$ If $a \preccurlyeq b$ and $b \preccurlyeq c$ then $a \preccurlyeq c$
3. $\forall\, a, b \in M :$ If $a \preccurlyeq b$ and $b \preccurlyeq a$ then $a = b$

We say $\preccurlyeq$ is a partial ordering on M.
A *lattice* is a partially ordered set in which every pair of elements (or, equivalently, every finite set of elements) has a supremum and an infimum. A *complete lattice* is a lattice in which every subset has a supremum (least upper bound) and an infimum (greatest lower bound). An *orthocomplementation* on a lattice M with partial ordering $\preccurlyeq$ is a function that maps each $a \in M$ to another element $a' \in M$ (its orthocomplement) such that

1. $(a')' = a$
2. $a' \cap a = 0$
3. $(a \preccurlyeq b)$ if and only if $(b' \preccurlyeq a')$

where $\cap$ denotes the *greatest lower bound* of a and a'.
The greatest lower bound of a *set* $\{a_i\}_{i \in J}$ where $a_i \in M$ for all $i \in J$, is defined as $\bigcap_J a_i$, where J is an index set. The greatest lower bound satisfies

$$x \preccurlyeq a_i \text{ for all } i \in J \text{ if and only if } x \preccurlyeq \bigcap_{i \in J} a_i.$$

[5]Note here that we cannot determine every observable precisely; Heisenberg's uncertainty principle prevents that.

The *least upper bound*, $\bigcup_J a_i = (\bigcap_J a_i')'$, has the property

$$a_i \preccurlyeq x \text{ if and only if } \bigcup_{i\in J} a_i \preccurlyeq x \quad \text{for all } x \in M.$$

This induces the notion of a zero element,

$$0 = \bigcap_{a\in M} a.$$

What exactly does our set of propositions look like, when we want to use it to represent quantum mechanical systems? The structure of a quantum mechanical system that matches empirical relations satisfies the following:

Postulate 1. *The propositions of a physical system form a complete and orthocomplemented lattice, denoted by* $\mathscr{L}$.

That is, $\mathscr{L}$ is partially ordered, every set of propositions has a greatest lower bound (which makes it both complete and a lattice) and it is orthocomplemented. Each element of $\mathscr{L}$ is a proposition (a yes-no experiment).

Now, we would like to interpret the concepts and relations we have seen earlier, and place this in the context of experimental questions and physical quantities. We read $a \preccurlyeq b$ (with $a, b \in \mathscr{L}$) as 'a implies b'. That is, b is true whenever a is true. If $a \preccurlyeq b'$, a and b are *disjoint*. In this 'proposition system'-context, the orthocomplementation is interpreted as negation: a' is the *negation* of a. For example, if a is the proposition 'The position of the particle is in the range $[0, 2]$', the proposition a' is 'The position of the particle is *not* in the range $[0, 2]$'. Lastly, we interpret the $\cap$ as conjunction. The proposition $a \cap b$ means 'a and b'.
The zero element of the lattice of propositions is the *absurd proposition*, $\emptyset = \bigcap_{a\in\mathscr{L}} a$, which is always false. We may now interpret the second property of an orthocomplementation as

$$a' \cap a = \emptyset.$$

Of course, when we have an absurd proposition, we also have the trivial proposition denoted by I: $I = \bigcup_{a\in\mathscr{L}} a$ is always true.

3.2 Measurement of States

In classical mechanics, it is possible to measure several quantities at the same time. One is therefore able to determine the 'state' the classical system is in; it is possible to attribute a value to each observable, and these values constitute the state of the system.
In quantum mechanics, however, a state can only be measured if the system can be prepared an unlimited number of times in the same state. Because of the inability to determine precisely the values of multiple observables simultaneously, we may not expect a definite value for each of the propositions at the same instant and thereby determine its state in a classical way. We may expect at most a probabilistic description of the observables. A quantum mechanical *state* is then mathematically determined by a real valued function $p : \mathscr{L} \to [0, 1], a \mapsto p(a)$, with the following properties:

1. $p(\emptyset) = 0$, $p(I) = 1$.
2. If $\{a_i\}$ is a countable and pairwise orthogonal sequence[6] then

$$\sum_i p(a_i) = p\left(\bigcup_i a_i\right).$$

[6] $\{a_i\} \subseteq \mathscr{L}$ is *pairwise orthogonal* if $a_i \preccurlyeq a_k'$ whenever $i \neq k$.

3. For any countable sequence $\{a_i\}$: if $p(a_i) = 1$ for all i, then $p\left(\bigcap_i a_i\right) = 1$

4. If $a \neq \emptyset$ there exists a p with $p(a) \neq 0$. If $a \neq b$ there exists a p such that $p(a) \neq p(b)$

so p is a *measure* on $\mathscr{L}$. Intuitively, a state p gives the probability $p(a)$ that a yes-no experiment a has a yes-outcome. Different states may give different probabilities to propositions. In the case of the EPR experiment, after measuring the first particle and concluding that its spin is in the up state, the proposition 'Is the spin of the first particle in the down state?' will be assigned the value 0. (Like all other propositions that do not ask whether the spin is 'up'.) Before measurement, this proposition will be assigned the value $\frac{1}{2}$, since the particle has probability $\frac{1}{2}$ of being in the up state.

The states form a convex set, that is: If p and q are two different states, then $\lambda p + \kappa q$ is again a state (where $\lambda, \kappa > 0$ and $\lambda + \kappa = 1$)[7]. If a state is represented as a combination of two others, we will call it a *mixed state*. If it cannot be represented as such, the state is called *pure*.

3.3 Compatibility

Let $M \subset \mathscr{L}, M \neq \emptyset$. M is a *sublattice* of $\mathscr{L}$ if it is a lattice closed under the operations of conjunction, disjunction and complementation; that is, for every $a, b \in M$ the expressions $a \cap b, a \cup b$ and a' are again elements of M.
We now introduce the important notion of compatibility for propositions: We may call two propositions x and y *compatible* when $S = \{x, y\}$ generates[8] a Boolean sublattice of $\mathscr{L}$. By a *Boolean (sub)lattice* $\mathscr{M}$ we mean that the (sub)lattice is distributive; every triple of elements of that lattice suffices the distributive law: For any $x, y, z \in \mathscr{M}$

$$x \cap (y \cup z) = (x \cap y) \cup (x \cap z).$$

We denote this compatibility between two propositions as $x \leftrightarrow y$, and we call the set of propositions that are compatible with all other propositions the *center* $\mathscr{C}$ of the lattice $\mathscr{L}$. Intuitively, the center of the lattice $\mathscr{L}$ consists of precisely those propositions that can be verified simultaneously. If the only propositions in the center are $\emptyset$ and I, the lattice has a *trivial center*.
We postulate:

Postulate 2 (Weak modular law)**.** *If $a \preccurlyeq b$, then $a \leftrightarrow b$.*

From this postulate it follows that if a and b are disjoint, then $a \leftrightarrow b'$.

We may define a *point* as a proposition $P \neq \emptyset$ that satisfies

1. $\emptyset \preccurlyeq P$

2. For every $a \in \mathscr{L}$ with $\emptyset \preccurlyeq a \preccurlyeq P$ we have either $a = \emptyset$ or $a = P$.

We now postulate the atomicity axiom:

Postulate 3. *Part I: For any proposition $a \neq \emptyset$, there exists a point P such that $P \preccurlyeq a$. (Existence of minimal propositions).*
Part II: If Q is a point and if we have

$$a \preccurlyeq x \preccurlyeq (a \cup Q), \quad \text{then} \quad x = a \text{ or } x = a \cup Q.$$

[7] See also Chapter 2.2.
[8] By a sublattice generated by a set S we mean that it is the smallest sublattice that contains all of S.

We redefine the notion of a *proposition system*: It is a system of elements $\mathscr{L}$ that is a complete, orthocomplemented lattice and satisfies postulates 2 and 3. As it happens, the weak modular law cannot be replaced by the *modular law*, as was proposed by the early investigators of the mathematics of quantum mechanics:

$$x \cup (y \cap z) = (x \cup y) \cap z \qquad \text{for all } x \preccurlyeq z,$$

because modularity and localizability are incompatible[9]. In other words, if the proposition would satisfy the modular law, one would know the probability for finding the system with the weak modular law without specifications about the preparation of the system. This doesn't reflect the physical reality; in real life, the method of preparation determines the outcome of the experiment. Since the best prediction of where for example a particle is, is a statistical distribution of the possible positions, the preparation determines the outcome of the experiment to a great extent. Therefore, the modular law was refuted and replaced by the weak modular law.

[9]For a proof, see Jauch, Chapter 12 [12]

4 Hilbert Space

The way to describe quantum mechanics is not by the lattice of propositions, but by self-adjoint operators on Hilbert space[10]. The lattice of propositions offers an intuitive view into the behaviour of elementary particles, but the lattice of Hilbert space makes the theory mathematically much more comprehensible and clarifies it, since Hilbert spaces are well understood. Before we can make this credible, we must first examine the Hilbert space itself.
A Hilbert space $\mathscr{H}$ is a vector space with an inner product that is complete with respect to the norm induced by its inner product. We express this as a set of axioms:
Axiom 1: $\mathscr{H}$ is a linear vector space over the complex numbers.
Axiom 2: There exists a strictly positive scalar product in $\mathscr{H}$, satisfying the following conditions:

$$\begin{aligned} (x,y) &= (y,x)^* \\ (x,y+z) &= (x,y)+(x,z) \\ (x,\lambda y) &= \lambda(x,y) \\ \|x\|^2 &= (x,x)\geq 0;\ \|x\|^2 = 0 \text{ if and only if } x=0. \end{aligned}$$

where $(x,y)^*$ is the complex conjugate of (x,y) and $x,y,z\in\mathscr{H}, \lambda\in\mathbb{C}$.
Axiom 3: $\mathscr{H}$ is complete: All Cauchy sequences in $\mathscr{H}$ converge.

Axiom 1 and 2 express that $\mathscr{H}$ is an inner product space, and Axiom 3 makes it complete. We will assume from now on that every Hilbert space is separable as well, and we add:
Axiom 4: $\mathscr{H}$ is separable, that is, there exists a sequence $\{x_n\}$ of vectors in $\mathscr{H}$ that is dense in $\mathscr{H}$: For any $x\in\mathscr{H}$ and for any $\varepsilon>0$ there exists at least one element x_n of the sequence such that

$$\|x-x_n\|<\varepsilon.$$

Note that both Axioms 3 and 4 hold automatically when the inner product space is finite-dimensional. If the inner product space is infinite-dimensional, these axioms have to be postulated, because they are not trivial for infinite-dimensional inner product spaces. Therefore we have added them as axioms.
For any sequence $\{x_n\}$ of linearly independent vectors in a Hilbert space, all vectors x that can be expressed as $x=\sum_n \lambda_n x_n$ together form a *linear manifold*. Any closed linear manifold in a Hilbert space is called a *subspace* of that Hilbert space. As we will see, propositions can be associated with subspaces of Hilbert space and observables are represented as self-adjoint (Hermitian) operators in a Hilbert space.
In any Hilbert space, every subspace M of that Hilbert space has a complement $M^\perp$ which is again in the Hilbert space, where $M^\perp=\{a\in\mathscr{H} : a\perp b \text{ for all } b\in M\}$.
A *projection* P is a linear operator that satisfies $P^2=P$. Let $\mathbb{P}$ be the space of projections on the subspaces of $\mathscr{H}$. We can associate with each subspace $M\subset\mathscr{H}$ a projection P^M on that subspace. This space $\mathbb{P}$ is a complete and orthocomplemented lattice if we interpret $M\vee N$ as $M+N$, the span of M and N, and $M\wedge N$ as $M\cap N$. The complement is as defined above, and the partial ordering is by set inclusion, $\subseteq$.

4.1 Observables and spectral measures

Each observable can be described using a *σ-homomorphism*: A mapping $x : B(\mathbb{R}) \to \mathscr{L}$, $\Delta \mapsto x(\Delta)$ (where $B(\mathbb{R})$ is the family of Borel sets on the real line) is a σ-homomorphism if it satisfies:

[10] As we will see later, the lattice of propositions is actually equivalent to the lattice of subspaces of Hilbert space.

1. $x(\mathbb{R}) = I$
2. $\Delta_1 \cap \Delta_2 = \emptyset$ implies that $x(\Delta_1) \perp x(\Delta_2)$
3. For any sequence $\{\Delta_i\}$ of pairwise disjoint Borel sets, we have

$$x\left(\bigcup_i \Delta_i\right) = \bigcup_i x(\Delta_i).$$

Note here that $x(\Delta_1) \perp x(\Delta_2)$ means that the propositions are disjoint, i.e. $x(\Delta_1) \preccurlyeq (x(\Delta_2))'$. Note furthermore that from 1 and 3 it follows that

$$I = x(\mathbb{R}) = x(\mathbb{R} \cup \emptyset) = x(\mathbb{R}) \cup x(\emptyset) = I \cup x(\emptyset),$$

so that $x(\emptyset) = \emptyset$. We may interpret this σ-homomorphism as an $\mathscr{L}$-valued measure. Examples of observables include 'spin', 'position', 'velocity' et cetera.

Since we can associate to each proposition a real value between 0 and 1 by means of state functions (see 3.2), we may introduce the function

$$\alpha_{p,x}(\Delta) = p(x(\Delta)),$$

which gives the probability of finding the value of x in the set Δ when the system is in state p. This explains why we used the σ-algebra of Borel sets as the domain for the observables; it contains every open and every closed interval, and measure theory works very well with σ-algebras. The composition of a state and an observable is again a measure. We may call this the *expectation value* of Δ.

It turns out that a σ-homomorphism is in fact a *spectral measure*. A spectral measure is a function $\mu : B(\mathbb{R}) \to \mathbb{P}$ that satisfies the following equations for all $\Delta, \Delta_1, \Delta_2 \in B(\mathbb{R})$;

1. $\mu(\Delta_1) + \mu(\Delta_2) = \mu(\Delta_1 \cup \Delta_2)$,
2. $\mu(\Delta_1) \cap \mu(\Delta_2) = \mu(\Delta_1 \cap \Delta_2)$,
3. $\mu(\emptyset) = 0$, where 0 is the projection into the null space ($0(\Delta) = 0$),
4. $\mu(\mathbb{R}) = I$, where I is the identity projection ($I(\Delta) = \Delta$).

We see that indeed, the conditions for the σ-homomorphism can be transformed into these conditions.

4.2 Compatibility

Two subspaces M and N of $\mathscr{H}$ are compatible, written as $M \leftrightarrow N$, when the lattice generated by M and N is a Boolean sublattice. That is, if it satisfies

$$(M \cap N) + (M \cap N^\perp) = N.$$

Here we mean by 'a lattice generated by M and N' (or, generally, by a set of subspaces) that this lattice is the smallest lattice containing both subspaces. In other words, if by $\mathscr{L}_i$, $i \in I$ we denote the sublattices that contain both M and N, the lattice generated by M and N is $\mathscr{L}_{M,N} = \bigcap_i \mathscr{L}_i$. Furthermore, we will call a lattice Boolean when every triple of subsets of $\mathscr{H}$ satisfies the distributive law.

In quantum mechanics, not all propositions are simultaneously measurable. Only propositions that are compatible with each other are simultaneously measurable. We may extend this notion of compatibility to observables. We will show that two observables are *compatible* if the Boolean sublattices of $\mathscr{L}$ generated by these observables are compatible. We will explain whaw we mean by this exactly.
The *range* of some observable x (which is now a σ-homomorphism) is the subset $\mathscr{B}_x \subset \mathscr{L}$ defined by

$$\mathscr{B}_x = \{x(\Delta) : \Delta \in B(\mathbb{R})\}.$$

This is a Boolean sublattice of $\mathscr{L}$ (since $B(\mathbb{R})$ is Boolean).
A Boolean sublattice $\mathscr{B}$ of $\mathscr{L}$ is called *separable* if there exists a countable subset $\{x_i\}$ of $\mathscr{B}$ such that $\mathscr{B} = \mathscr{B}(\{x_i\})$, that is, $\mathscr{B}$ is generated by this subset. (In other words, $\mathscr{B}(\{x_i\})$ is the smallest sublattice of $\mathscr{L}$ that contains $\{x_i\}$) We call $\mathscr{L}$ separable whenever every Boolean sublattice is separable. Furthermore, a Boolean sublattice $\mathscr{B}$ of $\mathscr{L}$ is separable if and only if there exists an observable x such that $\mathscr{B} = \mathscr{B}_x$. The proof of this claim is in Varadarajan, Lemma 3.16 [19]. We now arrive at the definition of compatibility:
Two observables x and y are *compatible* if every proposition of $\mathscr{B}_x$ is compatible with every proposition of $\mathscr{B}_y$.
For subspaces of a Hilbert space, we may express compatibility as follows: Let M and N be subspaces of a Hilbert space and P^M and P^N their corresponding projections, respectively. This correspondence means that to each projection we associate the subspace formbed by range of that projection. Thus, every subspace corresponds to a projection and vice versa. We have the following lemma:

Lemma 1. *Let $M_1, M_2 \in \mathscr{L}(\mathscr{H})$, where $\mathscr{L}(\mathscr{H})$ is the lattice of subspaces of $\mathscr{H}$ (see below). Then:*

$$M_1 \leftrightarrow M_2 \quad \text{if and only if} \quad P^{M_1} \text{ and } P^{M_2} \text{ commute,}$$

where P^M is the orthogonal projection onto the subspace $M \subseteq \mathscr{H}$.

Proof. Let $M_1 \leftrightarrow M_2$. Therefore, the lattice $\mathscr{L}_{M_1,M_2}$ is distributive. And so we have

$$M_1 = M_1 \cap \mathscr{H} = M_1 \cap (M_2 + M_2^\perp) = (M_1 \cap M_2) + (M_1 \cap M_2^\perp)$$

and

$$M_2 = (M_1 \cap M_2) + (M_1^\perp \cap M_2),$$

where '$+$' denotes the span of two subspaces. Now we define $N = (M_1 \cap M_2), N_1 = (M_1 \cap M_2^\perp)$ and $N_2 = (M_1^\perp \cap M_2)$. Note that N, N_1 and N_2 are mutually orthogonal. We write

$$M_1 = N_1 + N \quad \text{and} \quad M_2 = N_2 + N.$$

Then indeed, P^{M_1} and P^{M_2} commute:

$$\begin{aligned} P^{M_1}P^{M_2} &= (P^{N_1} + P^N)(P^{N_2} + P^N) \\ &= P^{N_1}P^{N_2} + P^{N_1}P^N + P^N P^{N_2} + P^N P^N \\ &= P^N. \end{aligned}$$

Because as N, N_1 and N_2 are mutually orthogonal, $P^N P^{N_2}$, $P^{N_1}P^{N_2}$ and $P^{N_1}P^N$ are all zero. And $P^N P^N$ is just P^N. In the same way we obtain

$$P^{M_1}P^{M_2} = P^N.$$

So indeed, P^{M_1} and P^{M_2} commute.

For the converse, assume that P^{M_1} and P^{M_2} commute. We write

$$Q_1 = P^{M_1} - P \quad \text{and} \quad Q_2 = P^{M_1} - P.$$

Then Q_1 and Q_2 are again projections and

$$PQ_1 = P(P^{M_1} - P) = P - P^2 = P - P = 0.$$

In the same way it follows that $Q_1P = PQ_2 = Q_2P = Q_1Q_2 = Q_2Q_1 = 0$. So now Q_1, Q_2 and P are mutually orthogonal. If N_1, N_2 and N are the closed manifold ranges corresponding to the projections Q_1, Q_2 and P respectively, it follows that these are mutually orthogonal, too. In addition, we have $M_1 = N_1 + N$ and $M_2 = N_2 + N$. This proves that $M_1 \leftrightarrow M_2$. □

We will call a set $\mathscr{L}(\mathscr{H})$ of *subspaces of* $\mathscr{H}$ a lattice of subspaces (or simply: lattice) when $\mathscr{L}(\mathscr{H})$ is closed with respect to the operations $\cap$ and $+$. Here $+$ denotes the span of two subspaces and $\cap$ is the intersection of the two subspaces.
Since $\mathscr{L}(\mathscr{H})$ is closed with respect to a countably infinite number of intersections and spans, and for each $M \in \mathscr{L}(\mathscr{H})$ there exists a $M^\perp \in \mathscr{L}(\mathscr{H})$, we obtain a complete, orthocomplemented lattice. As we have postulated in Chapter 3.1, the propositions of a physical system are such a complete, orthocomplemented lattice. How is the lattice of propositions of a physical system related to the lattice of subspaces of Hilbert space? It turns out that there is a close relation between the two. We examine this relation in the next chapter.

5 The Connection Between the Lattice of Propositions and Hilbert Space

In this section, we will try to make clear how we connect our lattice of propositions from Section 3 to the Hilbert spaces we saw in Section 4. As it happens, this connection leads us via the projective geometries, which we will introduce here. After that, with the aid of a helpful theorem, we will close the gap between propositions and subspaces of Hilbert space.

5.1 Projective Geometries

Why is using subspaces of a Hilbert space the way of describing quantum mechanics? To be frank, there is not a clear physical answer to this question yet, but the bridge between an abstract proposition and a Hilbert space can be constructed using projective geometries. We will give a brief explanation of projective geometries, in order to understand what this 'bridge' does, exactly.
A projective geometry is defined as a class G of subsets of E, where E is a set of elements called points. These subsets of E satisfy the following properties:

1. There exists a class of subsets called 'lines' in G such that for each pair of different points e_1, e_2 there exists exactly one line l that contains both e_1 and e_2.

2. Three points e_1, e_2 and e_3 that are not all on the same line form a triangle that has the following property: If e_{12} is a point on the line through e_1 and e_2, and e_{23} is a point on the line through e_2 and e_3, then e_{12} and e_{23} determine a line l that contains a point e_{13} on the line through e_1 and e_3.

3. The necessary and sufficient condition that a subset of points belongs to the class of subsets G is that it contains the lines that pass through any pair of points from that subset.

If the partial ordering in G is set inclusion, and unions and intersections are defined as least upper bound and greatest lower bound respectively, then G is an atomic modular lattice.
It is important to note here that parallel lines or planes do not exist in projective geometries. This can be deduced from axiom 2. If we take points e_{12} and e_{23} as in Figure 1 there should still be a point where the two lines meet, according to axiom 2. We take this point to be at infinity (both negative and positive infinity, which are taken as the same here).

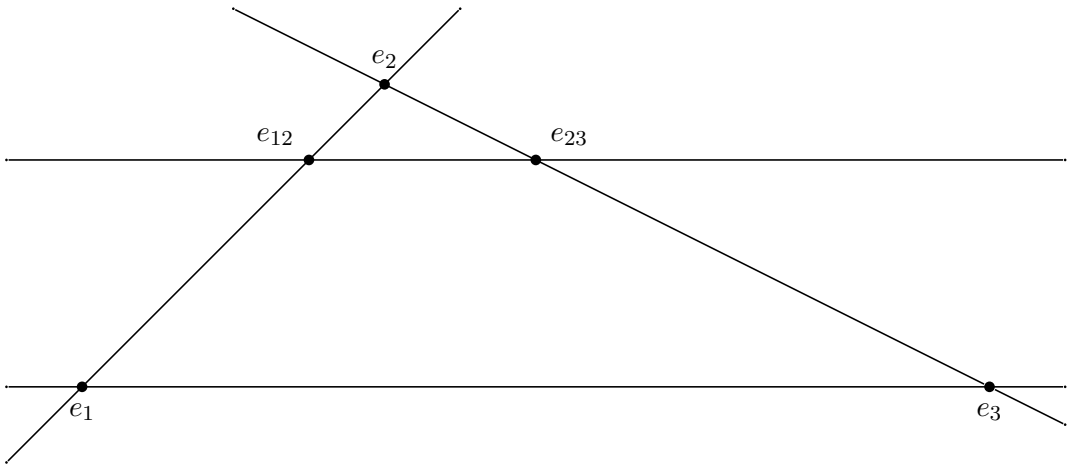

Figure 1: Parallel lines meet at infinity

The dimension of a projective geometry is determined by the existence of independent sets. If there are, for example, two distinct points in E, the dimension is *at least* 1. The dimension of the geometry is *at most* 1 when there is at most one line in the projective space.

5.2 Reducible Lattices

It is known[11] that the properties of projective geometries are expressible in terms of intersections and unions of geometrical elements. The structure of a projective geometry is that of an atomic lattice. This seems a very nice result at first. This suggests that we could easily connect a projective geometry to our lattice. However, a projective geometry is *modular*, whereas our proposition system only satisfies the *weak* modular law (see Section 3.3). We will therefore explain the notion of reducible lattices. Reducible lattices help us to associate projective geometries to our lattice of propositions.
Let $\mathscr{L}$ be a lattice and $\mathscr{C}$ its center. That is, if $a \in \mathscr{C}$, then a is compatible with every proposition in $\mathscr{L}$. If $\mathscr{C} = \{\emptyset, I\}$, the center is trivial. In that case, we will call the lattice *irreducible*. If the center is not trivial, we can reduce the lattice in the following way.
If $\mathscr{C}$ is not trivial, there must be at least one element a that is an element of $\mathscr{C} \setminus \{\emptyset, I\}$, and thus also the complement, $a' \in \mathscr{C} \setminus \{\emptyset, I\}$. This follows from the axioms of a proposition system, see Chapter 3.1. We may then write for each element $x \in \mathscr{L}$:

$$x_1 = x \cap a \quad \text{and} \quad x_2 = x \cap a'.$$

Both a and a' are elements of $\mathscr{C}$, so the distributive law holds for the expression

$$x \cap (a \cup a') = (x \cap a) \cup (x \cap a'),$$

and it follows that

$$x_1 \cap x_2 = (x \cap a) \cap (x \cap a') = x \cap (a \cap a') = x \cap \emptyset = \emptyset.$$

So x_1 and x_2 are disjoint propositions. We may now write x as an ordered pair (x_1, x_2). If $x, y \in \mathscr{L}$ are different propositions, then x and y can be written respectively as $x = (x_1, x_2)$ and $y = (y_1, y_2)$. We have

$$x \cap y = (x_1 \cap y_1, x_2 \cap y_2) \quad \text{and} \quad x \cup y = (x_1 \cup y_1, x_2 \cup y_2).$$

The orthocomplement is $x' = (x_1', x_2')$, where $x_1' = x' \cap a$ and $x_2' = x' \cap a'$. This completes the reduction of $\mathscr{L}$ into two sublattices $\mathscr{L}_1 = \{x \cap a : x \in \mathscr{L}\}$ and $\mathscr{L}_2 = \{x \cap a' : x \in \mathscr{L}\}$.
Of course, $\mathscr{L}_1$ and $\mathscr{L}_2$ both contain a center, $\mathscr{C}_1$ and $\mathscr{C}_2$, which may be trivial or not. If one of them is or both are not trivial, we can reduce these lattices even further. If such a center is trivial, the decomposition stops.

Even if the lattice is non-Boolean or infinitely large, such a decomposition can be made. Let $\mathscr{L}$ be a non-Boolean lattice with a non-trivial center. Let P the collection of points. We will call points e_1 and e_2 in *perspective* if there exists an $e_3 \in P$ with

$$e_3 \preccurlyeq e_1 \cup e_2.$$

We note this as $e_1 \sim e_2$. Perspectivity is an equivalence relation. Indeed, $e_1 \sim e_1$ since $e_1 \preccurlyeq e_1 \cup e_1 = e_1$ (see the first requirement of the partial ordering in $\mathscr{L}$). Furthermore, if there exists an e_3 such that $e_3 \preccurlyeq e_1 \cup e_2$, then that e_3 also satisfies $e_3 \preccurlyeq e_2 \cup e_1$ and thus perspectivity is symmetric. For transitivity, we have Lemma 2, where we already adopt the terminology we used for geometries: By a line determined by two points e_1 and e_2 (that are perspective) we mean the set $\{e : e \preccurlyeq e_1 \cup e_2\}$.

[11] See Jauch, Chapter 8.1 [12]

Lemma 2. *Let e_1, e_2 and e_3 be three distinct points that are not on the same line. If e_4 is on the line $e_1 \cup e_2$ and different from both e_1 and e_2, and e_5 is a point on the line $e_2 \cup e_3$ different from both e_2 and e_3, there exists a line $l = e_4 \cup e_5$. This line contains exactly one point, e_6 on the line $e_1 \cup e_3$, different from e_3 and e_1.*

For the proof, see Piron [15]. If we have such an $e_4 \preccurlyeq e_1 \cup e_2$ and $e_5 \preccurlyeq e_2 \cup e_3$, then e_6 is the point that we need for e_1 to be perspective to e_3: $e_6 \preccurlyeq e_1 \cup e_3$.

We can use this perspectivity to divide $\mathscr{L}$ into equivalence classes. Let $[e]$ be the class of points perspective to the point $e \in \mathscr{L}$. We denote by z_e the collection

$$z_e = \bigcup_{a \in [e]} a.$$

Of course, we have $\bigcup\{z_e : [e] \text{ is an equivalence class } \} = I$. Each such z_e is contained in the center $\mathscr{C}$ of $\mathscr{L}$ and it is the smallest element in the center that contains all points in $[e]$. We call z_e the *central cover of the points in* $[e]$. Now, every element $x \in \mathscr{L}$ can be written as the union of points $x = \bigcup_i x_i$, $i \in \mathscr{L}$ where $x_i = x \cap z_i$. These elements x_i form an irreducible lattice closed under union and intersection: if $x = \bigcup_i x_i$ and $y = \bigcup_i y_i$, then $x \cup y = \bigcup_i (x_i \cup y_i)$ and $x \cap y = \bigcap_i (x_i \cap y_i)$. This ensures that every lattice can be reduced to irreducible sublattice, regardless of whether $\mathscr{L}$ is infinitely large or whether it is Boolean or not. For the details of this construction, confer Piron [16].

But despite this elaborate construction, the question remains: Why would we want such a decomposition? The reduction of the lattice $\mathscr{L}$ into two lattices $\mathscr{L}_1$ and $\mathscr{L}_2$ is very helpful, because every irreducible proposition system can be imbedded into a projective geometry:

Theorem 1. *Every irreducible proposition system $\mathscr{L}$ can be imbedded in a canonical way into a projective geometry G by a function $\alpha : \mathscr{L} \to G$ that has the following properties:*

1. *The restriction of α to the points of $\mathscr{L}$ is a mapping into the points of G*
2. *$a \preccurlyeq b$ if and only if $\alpha(a) \subseteq \alpha(b)$*
3. *$\alpha\left(\bigcap_i a_i\right) = \bigcap_i \alpha(a_i)$*
4. *$\alpha(a \cup e) = \alpha(a) \cup \alpha(e)$ for any point $e \in \mathscr{L}$*

Before we continue, we define the concept of a *line* in a proposition system: For points p and r, the line through both of them is denoted as $pr := \{q : q \text{ is a point and } q \preccurlyeq p \cup r\}$.

Proof. Let $\alpha : \mathscr{L} \to G$ be the function that associates with each proposition $a \in \mathscr{L}$ the set of points e with the property $e \preccurlyeq a$: $\alpha(a) = \{e : e \in P(\mathscr{L}) \text{ and } e \preccurlyeq a\}$. We will show that this function has all four properties.
Note, first, that by $\alpha(a \cup e)$ we mean the set $\{p : p \text{ is a point of } \mathscr{L} \text{ and } p \preccurlyeq a \cup e\}$ and by $\alpha(a) \cup \alpha(e)$ we mean the supremum of $\alpha(a)$ and $\alpha(e)$: it is the union of all lines $p \cup e$ with $p \preccurlyeq a$.
Property 1 and 2 follow directly from our definition of α. For property 3, we analyze what this equality means. The left-hand-side consists of all points contained in $\bigcap_i a_i$. The right-hand-side consists of all points contained in each of the a_i, so that is exactly $\alpha\left(\bigcap_i a_i\right)$.
For property 4, we have to do a little more work. First we have a Lemma that will help us:

Lemma 3. *For every variety a and every point e the following holds: If q lies in $a \cup e$, then q lies on a line er with r on a.*

By a *variety* we mean a subset of the set of lines E that contains the line eq whenever it contains the two (distinct) points e and q.

Proof of Lemma 3. Let $R = \{q : q \text{ lies on a line } er \text{ with } r \text{ on } a\}$. We prove that $R = a \cup e$. It is clear that $R \subseteq a \cup e$. We only have to prove that R is a variety; R would then be a variety that contains a and e, and $a \cup e$ is the smallest variety that contains a and e, so we would have $a \cup e \subseteq R$. In other words, we have to prove that if $q \preccurlyeq a \cup e$, then there exists an r on a such that $q \preccurlyeq e \cup r$, that is, q lies on the line er. Let $q \preccurlyeq a \cup e$ and assume that for some r on a, $r \preccurlyeq e \cup q$. Then $e \preccurlyeq e \cup r \preccurlyeq e \cup q$, and because of Postulate 3, $e = e \cup r$ or $e \cup r = e \cup q$. The first possibility can be disregarded since e is a point and $e \cup r$ clearly is not. Then it follows that q lies on er, since er and eq are the same line.

Now, we still have to prove that if $q \preccurlyeq a \cup e$, then there exists an r with $r \preccurlyeq q \cup e$. It suffices to prove $(e \cup q) \cap a \neq \emptyset$. Then, indeed, there would exist an r on a such that $r \preccurlyeq e \cup q$. Suppose that $(e \cup q) \cap a = \emptyset$. Then, $((e \cup q) \cap a))' = I$, which is the same as $(e \cup q)' \cup a' = I$.

Claim 1. *There exists an atom r such that*

$$(e \cup q)' \cup r = (e \cup q)' \cup a'.$$

Proof of Claim. We know that $a \cup (e \cup q) = a \cup e$ (because $q \preccurlyeq a \cup e$). We also have

$$\begin{aligned} a' &= (a' \cap e') \cup e \\ &= (a \cup e)' \cup e \\ &= (a \cup e \cup q)' \cup e \\ &= (a' \cap (e \cup q)') \cup e \end{aligned}$$

So there is an atom, namely $r = e$, that satisfies $a' = (a' \cap (e \cup q)') \cup r$. Thus

$$(e \cup q)' \cup a' = (e \cup q)' \cup (a' \cap (e \cup q)') \cup r = (e \cup q)' \cup r,$$

and therefore, $(e \cup q)' \cup a' = (e \cup q)' \cup r$. □

So we have $(e \cup q)' \cup r = I$. We also have the weak modular law:

$$(e \cup q) \cap ((e \cup q)' \cup r)) = r$$

Using the fact that $(e \cup q)' \cup r = I$, we get

$$(e \cup q) \cap I = r,$$

but of course, we also have $(e \cup q) \cap I = (e \cup q)$. This contradicts our assumption that $e \cup q$ is not a point. So $(e \cup q) \cap a \neq \emptyset$. □

So indeed, for our projective geometry, $\alpha(a \cup e) = \alpha(a) \cup \alpha(e)$. And so all properties of the function $\alpha : \mathscr{L} \to G$ are proven. □

Our result is completed by the following theorem:

Lemma 4. *If G is a projective projective geometry, then Lemma* 3 *holds.*

The proof is in Piron, [16].

What we have shown, or at least made plausible, is that every proposition system is a direct union of irreducible proposition systems, and that those irreducible proposition systems can be imbedded into projective geometries. The next step is to associate with the projective geometries the subspaces of Hilbert space.

5.3 Closing the Gaps

The association between projective geometries and subspaces of Hilbert space uses the following theorem:

Theorem 2. *If a projective geometry G has dimension $n \geq 3$, then there exists a linear vector space V with coefficients from a field F and a bijection between the elements of G and the linear manifolds of V. This bijection preserves the partial ordering and maps points to one-dimensional subspaces of V.*

The proof can be found in Baer, Chapter VII [1]. It is somewhat too advanced to include it here.
This bijection is the last piece of the bridge between propositions and subspaces of Hilbert space. As mentioned before, there is no physical basis for choosing Hilbert spaces. It has been shown[12] that the only options for the field of coefficients are the Banach division algebras $\mathbb{R}$, $\mathbb{C}$ and the quaternions.
However, the cases of real and quaternionic Hilbert space can be reduced to that of complex Hilbert space, or at least for simple systems (See Jauch, Chapter 8 [12]).
Although there is no physical basis for choosing Hilbert spaces, there is some motivation behind choosing this particular kind of vector space. First we introduce the concept of a C^*-algebra:
A *C^*-algebra* is a (complex) Banach space X, together with a product $X \times X \to X$ that sends every pair of elements $x, y \in X$ to $x \cdot y$. This product is continuous, and in a Banach algebra we have $\|x \cdot y\| \leq \|x\|\|y\|$ for all $x, y \in X$. Additionaly, the Banach space contains a unit element $1 \in X$ with $1 \cdot x = x \cdot 1 = x$ for all $x \in X$. Moreover, it contains an involution $^* : x \mapsto x^*$ that satisfies the following:

1. $x^{**} = x$,
2. $(\lambda x)^* = \overline{\lambda} x^*$, where $\overline{\lambda}$ is the complex conjugate of λ,
3. $(x + y)^* = x^* + y^*$,
4. $(x \cdot y)^* = y^* \cdot x^*$, and
5. $\|x^* \cdot x\| \leq \|x^*\|\|x\|$

for all $x, y \in X$.
Now, we have the following theorem:

Theorem 3 (Gelfand, Naimark)**.** *Any C^*-algebra X is isometrically *-isomorphic to a C^*-algebra of bounded operators on a Hilbert space.*

Here, 'isometrically' means distance-preserving, and '*-isomorphic' means that the isomorphism respects the involution. The proof can be found in Kadison [13] or in their original paper, [8].

The C^*-algebra is a generalization of the commutative C^*-algebras used in classical mechanics. In this context, observables are self-adjoint elements of the C^*-algebra (that we will call $\tilde{A}$). By the Gelfand-Naimark theorem, these are the self-adjoint bounded operators on Hilbert space. The states are positive functionals $\varphi : \tilde{A} \to \mathbb{C}$, with $\varphi(u^*u) \geq 0$ for all $u \in \tilde{A}$ and $\varphi(1) = 1$. So the expectation value of an observable x in state φ is then $\varphi(x)$.

We will now assume that the vector space is Hilbert space, and that the field of coefficients is $\mathbb{C}$. Elementary propositions are now represented by projection operators or, equivalently, by subspaces of Hilbert space, since we can associate with each proposition the orthogonal

[12] This is the Gelfand-Mazur theorem. See [14], Proposition 5.7 and 5.8.

projection on the corresponding subspace. Intersection and union are defined as set intersection and the span of subspaces.
For each proposition, that is for each subspace, there is a corresponding projection. Two propositions are *compatible* if and only if their corresponding projections commute (see Section 4.2). We have already mentioned that observables are represented by σ-homomorphisms $B(\mathbb{R}) \to \mathscr{L}$, which are now nothing else than spectral measures (see also 4.1). But spectral measures are in one-to-one correspondence with the self-adjoint linear operators:
First, we assume that the Hilbert space is finite-dimensional. The statement can be extended to the infinite-dimensional case, but we will not do that here[13]. Let Λ be the set of all eigenvalues of the operator associated with the observable A, and let $\lambda_r \in \Lambda$. The set $\{\psi_i\}$ for which

$$(A - \lambda_r I)\psi_i = 0$$

holds, spans a finite-dimensional space M_r, the eigenspace of A associated with eigenvalue λ_r. Let $\Delta \in B(\mathbb{R})$. To each such Δ we may associate a projection

$$E(\Delta) = \sum_{\lambda_r \in \Delta} P^{M_r},$$

where P^{M_r} is the projection on the subspace M_r. Then E is a spectral measure. This leads to the definition that observables are self-adjoint linear operators.
We can also write an observable A as a linear combination of pairwise orthogonal projections.

Theorem 4 (Spectral theorem)**.** *Let A be an observable and let M_r be again the eigenspace corresponding to the eigenvalue λ_r of A. Let P^{M_r} be the projection onto M_r. Then A can be decomposed as*

$$A = \lambda_1 P^{M_1} + \ldots + \lambda_n P^{M_n}.$$

Since we assumed the Hilbert space was finite-dimensional, we may view operators as matrices. The number of eigenvalues of A is the same as the magnitude of the dimension of $\mathscr{H}$, and eigenvectors corresponding to different eigenvalues are orthogonal. A is then orthogonally diagonalizable: $A = PDP^*$, where the columns of P are the orthonormal eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of A, P^* is the transpose of P and D is the diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ on the diagonal. Our observable A can now be written as

$$A = \lambda_1 \mathbf{u}_1 \overline{\mathbf{u}_1}^T + \ldots + \lambda_n \mathbf{u}_n \overline{\mathbf{u}_n}^T.$$

Every term in this expression is a matrix of rank 1 (since the eigenvectors are orthonormal). Furthermore, we can view $\mathbf{u}_i \overline{\mathbf{u}_i}^T$ as the orthogonal projection onto the one-dimensional subspace spanned by $\mathbf{u}_i$. We show that for every vector $\mathbf{x}$ in the subspace M_i there is an orthogonal projection of $\mathbf{x} = (x_1\ x_2\ \ldots\ x_n)^T$ onto a subspace M_i of $\mathscr{H}$ of the form

$$P^{M_i} = \mathbf{u}_i \overline{\mathbf{u}_i}^T \mathbf{x}.$$

The orthogonal decomposition theorem from Linear Algebra states that $\mathbf{x}$ can be written as

$$\mathbf{x} = \frac{(\mathbf{x}, \mathbf{u}_i)}{(\mathbf{u}_i, \mathbf{u}_i)} \mathbf{u}_i.$$

Since all eigenvectors are orthonormal, this is just

$$\mathbf{x} = (\mathbf{x}, \mathbf{u}_i)\mathbf{u}_i$$

[13] For the extension to the infinite-dimensional case, see Chapter 4.6 of Jauch [12]

If we write $\mathbf{u}_i = (u_{1,i}\ u_{2,i}\ \dots\ u_{n,i})^T$, we write the former as

$$\begin{aligned}
\mathbf{x} &= \begin{pmatrix} (\mathbf{x}, \mathbf{u}_i)u_{1,i} \\ \vdots \\ (\mathbf{x}, \mathbf{u}_i)u_{n,i} \end{pmatrix} \\
&= \begin{pmatrix} u_{1,i} \\ \vdots \\ u_{n,i} \end{pmatrix} (\mathbf{x}, \mathbf{u}_i) \\
&= \begin{pmatrix} u_{1,i} \\ \vdots \\ u_{n,i} \end{pmatrix} \left[\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} \overline{u_{1,i}} \\ \vdots \\ \overline{u_{n,i}} \end{pmatrix} \right] \\
&= \begin{pmatrix} u_{1,i} \\ \vdots \\ u_{n,i} \end{pmatrix} \begin{pmatrix} \overline{u_{1,i}} & \dots & \overline{u_{n,i}} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.
\end{aligned}$$

The product of the first two vectors is the projection matrix on the line through $\mathbf{u}_i$. So P^{M_i} may be written as $\mathbf{u}_i \overline{\mathbf{u}_i}^T$. Using the spectral measure from above, we may now write A as

$$A = \sum_{i=1}^{n} \lambda_i P^{M_i} = \int_{\mathbb{R}} \lambda \mathrm{d}E(\lambda).$$

This decomposition can be extended to the case where $\mathscr{H}$ is infinite-dimensional.

Theorem 5 (Spectral theorem - infinite case)**.** *To each self-adjoint operator A corresponds a unique projection-valued measure P_A such that*

$$A = \int_{\mathbb{R}} \lambda \mathrm{d}P_A(\lambda).$$

A projection-valued measure is a measure on $B(\mathbb{R})$. It is a family of projections $\{P^\Delta\}$, $\Delta \in B(\mathbb{R})$, that obeys the following

1. P^Δ is an orthogonal projection, where $\Delta \in B(\mathbb{R})$,
2. $P^\emptyset = 0$ and $P^{(-a,a)} = I$ for some $a \in \mathbb{R}$,
3. If $\Delta = \bigcup_{i=1}^{\infty} \Delta_i$ with $\Delta_i \cap \Delta_j = \emptyset$ whenever $i \neq j$, then

$$P^\Delta = \lim_{N \to \infty} \sum_{i=1}^{N} P^{\Delta_i},$$

4. $P^{\Delta_1} P^{\Delta_2} = P^{\Delta_1 \cap \Delta_2}$ for all $\Delta_1, \Delta_2 \in B(\mathbb{R})$.

In particular, there is a one-one correspondence between self-adjoint operators A and projection-valued measures:

$$A \mapsto \{P^\Delta\} = \{\chi_\Delta(A)\}, \quad \text{and} \quad \{P^\Delta\} \mapsto A = \int \lambda \mathrm{d}P(\lambda)$$

where $\Delta \in B(\mathbb{R}), \lambda \in \Lambda$ and $\chi_\Delta(A)$ is the characteristic function of A on the interval Δ.

States are functionals on the projection operators. An important theorem by Gleason says:

Theorem 6 (Gleason's Theorem)**.** *Each state p, that is each functional on the projection operators, has the form*

$$p(E) = \mathrm{Tr}\,(WE),$$

where Tr *denotes the trace function, W is a self-adjoint positive semi-definite operator (called the density operator) which satisfies*

$$\mathrm{Tr}\,W = 1$$

and E is such a projection operator (proposition).

The proof of this theorem may be found in Gleason [9].
We may now define the *expectation value* of an observable A by

$$p(A) = \mathrm{Tr}\,(WA).$$

As we identified observables with self-adjoint linear operators, this operation is well-defined.

We have seen that it is possible to see the lattice of propositions, viewed as yes-no experiments, as a lattice of subspaces of Hilbert space. Every concept in the lattice of propositions has its counterpart in $\mathscr{L}(\mathscr{H})$. We can summarize these results in a table:

	Lattice of propositions	**Hilbert space**
Partial ordering	$\preccurlyeq$	$\subseteq$
Proposition	(A, Δ)	subspace of $\mathscr{H}$
State	measure on $\mathscr{L}$	functional on projection operators
Observable	σ-homomorphism	spectral measures/self-adjoint linear operators
Expectation value	$p(x(\Delta))$	$p(A) = \mathrm{Tr}\,(WA)$

6 Superposition and Entanglement

Superposition and Entanglement both play a large role of the EPR paradox. Therefore, it would be nice if we could examine these concepts a little closer, and describe these phenomena in a mathematical way in order to be able to speak about them in a clear way.

6.1 Superposition

A characteristic trait of quantum mechanics is the *superposition* of states. If a physical system can be in a certain number of configurations, then the system will always be in all of these possible configurations simultaneously. For example, if the system could be in two configurations, 0 and 1, then the general description of the state the system is in is given by

$$z_1|0\rangle + z_2|1\rangle,$$

with $z_1, z_2 \in \mathbb{C}$ and $z_1^2 + z_2^2 = 1$. Here, $|\cdot\rangle$ denotes the Dirac bra-ket notation.

We can also give a mathematical account of what superposition means. Recall that a *point* is a proposition $e \neq \emptyset$ with the property that if $x \preccurlyeq e$ then $x = \emptyset$ or $x = e$.
Superposition may be defined as follows: If e_1, e_2 and are two different points, then there exists a third, e_3 such that $e_3 \neq e_1$ and $e_3 \neq e_2$ and

$$e_1 \cup e_2 = e_1 \cup e_3 = e_2 \cup e_3.$$

e_3 here plays the role of the superposition of e_1 and e_2. Note that a system in which the superposition principle is valid cannot be Boolean. For example, if we would have

$$(e_1 \cup e_2) \cup e_3 = (e_1 \cup e_3) \cup (e_2 \cup e_3)$$

Then the first expression between brackets would be equal to $e_1 \cup e_3$ (definition of e_3) and so the left-hand-side would be equal to $e_1 \cup e_3$. But then, on the other hand, if we examine the right-hand-side, $e_2 \cup e_3$ would be equal to $\emptyset$, in contradiction with the assumption that $e_1, e_2, e_3 \neq \emptyset$. Furthermore, e_3 cannot be compatible with both e_1 and e_2. (Otherwise the principle of distributivity would hold and the system would be Boolean.) In fact, we have the following theorem:

Theorem 7. *A proposition system which satisfies the principle of superposition for all pairs of points has trivial center.*

Recall that having a trivial center means that the only propositions that are compatible with all other propositions are $\emptyset$ and I. This theorem expresses the fact that superposition is not an axiom of quantum mechanics but merely a corollary of the non-Boolean structure of our proposition system. Jauch discusses the details of this theorem and its proof in Chapter 6.8[12], but for our goal, this is not of great importance.

6.2 Entanglement

As we have seen in Section 2.3, entanglement is the principle that the state probabilities of multiple particles can be dependent of each other. We have examples of two particles being entangled, but entanglement can take place between multiple particles. We have seen that the singlet state, the system of an entangled positron and an electron, is expressed as

$$\Psi = \frac{1}{\sqrt{2}} \left[(|\varphi_+\rangle \otimes |\psi_-\rangle) + (|\varphi_-\rangle \otimes |\psi_+\rangle \right]. \tag{2}$$

The fact that this state is a sum of two states, dependent of each other, means that the states are entangled: The probabilities of one state are dependent of the other state.

By contrast, consider a system of two particles, where particle 1 is in the state ψ_a and particle 2 is in the state ψ_b. Then the state that corresponds with the total system is

$$\psi = |\psi_a\rangle \otimes |\psi_b\rangle.$$

Here, we can distinguish which state belongs to which particle. For example, ψ_a would belong to the positron and ψ_b would belong to the electron. But trying to distinguish the states for the expression in Equation (2) would of course be problematic.

7 Quantum Logic

Can propositions of a quantum system be interpreted as sentences of a logical language? What rules does this language inherit from the ordered structure of propositions? In this section, we will try to answer these questions.

7.1 Sentences and Connectives

To the pair formed by the state p and the proposition a we can associate a sentence:

$$(p, a) \text{ means: "The physical system in state } p \text{ causes the yes-outcome of proposition } a\text{".} \quad (3)$$

Propositions generate predicates of the object language, while states play the role of individuals to which predicates apply. The physical system is assumed to be fixed here. For the sake of simplicity, we will only consider pure states.
The value of a at p is the probability of a yes-outcome of a when the system is in state p. However, here we encounter a problem, since with quantum mechanical systems (because of superposition) the probability of a yes-outcome for proposition a lies in the interval $[0, 1]$ and the logical truth value of a should be in $\{0, 1\}$. We therefore define (p, a) to be true whenever a gives a yes-outcome with probability 1. (p, a) is false in all other cases, that is, $p(a) \neq 1$. (This is not the only option we have. We could, for instance, allow the truth value to be in $[0, 1]$ (many-valued or fuzzy logic), but it turns out that this is not a very fruitful way to define the truth values for quantum logic. See also [3].) Note that this is in a sense (the classical sense) problematic: When we call a sentence (p, a) false, it is not necessarily so that (p, a) gives a no-outcome with certainty. It is just that the yes-outcome is not certain (but it is not impossible). We shall see that this problem resolves when we define the connectives.
First, let $\mathcal{S}(a)$ be the set of states that generate a yes-outcome of proposition a:

$$\mathcal{S}(a) = \{p : p(a) = 1\}.$$

This definition makes the mathematics more compact and the notation will be more friendly. Keep in mind that all states are *pure* states.

Now, how may we define these logical connectives? We will begin with the *conjunction*. Let (p, a) and (p, b) be two sentences. Since

$$\mathcal{S}(a \cap b) = \mathcal{S}(a) \cap \mathcal{S}(b),$$

this means that the sentence $(p, a \cap b)$ is true precisely whenever both (p, a) and (p, b) are true. The conjunction, $\wedge$, can now be defined as

$$(p, a) \wedge (p, b) \quad = \quad (p, a \cap b).$$

This is in accordance with the definition of conjunction in classical logic. Moreover, we write $a \wedge b$ for $\inf_p(p, a \cap b)$.
The *disjunction*, $\vee$, is defined as

$$(p, a) \vee (p, b) \quad = \quad (p, a \cup b).$$

Because in general, one has

$$\mathcal{S}(a) \cup \mathcal{S}(b) \subseteq \mathcal{S}(a \cup b),$$

there are states for which is $a \cup b$ is true but neither a nor b is. In fact, this is a feature of quantum mechanics. It is interesting to see that quantum disjunction has nonclassical features on the semantic level. In the same way as for conjunction, we write $a \vee b$ for $\inf_p(p, a \cup b)$.

For the *negation*, $\neg$, note that $\mathcal{S}(a')$ coincides with the certainly-no domain of (p, a). The negation may be defined as

$$\neg(p, a) \quad = \quad (p, a').$$

Again, there are non-classical features that make this definition a bit more complicated. The truth of (p, a) implies the falsehood of $\neg(p, a)$, but the falsehood of (p, a) *does not* imply the truth of $\neg(p, a)$, because of the asymmetric definiton of truth we adopted. This also means that there are states that make (p, a) false as well as $\neg(p, a)$. Furthermore, we write $\neg a$ for $\inf_p(p, a')$. We also have, by definition of disjunction, that

$$((p, a) \vee \neg(p, a)) \text{ is true for every } a \in \mathscr{L} \text{ and for each } p \in \mathcal{S}(a).$$

for every $a \in \mathscr{L}$. Note that, because there are states that make both (p, a) and $\neg(p, a)$ false, this does not mean that either (p, a) or $\neg(p, a)$ is true.
We also have the laws of De Morgan:

$$\begin{aligned} \neg((p, a) \vee (p, b)) &= \neg(p, a) \wedge \neg(p, b), \\ \neg((p, a) \wedge (p, b)) &= \neg(p, a) \vee \neg(p, b). \end{aligned}$$

Now, we would like to express the conditional connective in quantum logical notation. The conditional, noted as $\Rightarrow$, expresses 'If (p, a), then (p, b)', where (p, a) and (p, b) are two sentences. This is noted as $(p, a) \Rightarrow (p, b)$. One might try the classical conditional:

$$(p, a) \Rightarrow (p, b) \quad = \quad (p, a' \cup b),$$

but this only holds when $\mathscr{L}$ is Boolean, which is not the case. As it turns out, the correct choice is

$$(p, a) \Rightarrow (p, b) \quad = \quad (p, a' \cup (a \cap b)).$$

Note that this equals the classical conditional $a' \cup b$ when the lattice is Boolean. Furthermore, this conditional satisfies

$$(p, a) \Rightarrow (p, b) \text{ is true for every } p \quad \text{ if and only if } \quad a \preccurlyeq b,$$

and also the Modus Ponens rule,

$$[(p, a) \text{ and } ((p, a) \Rightarrow (p, b))] \Rightarrow (p, b) \text{ is true for every } p.$$

Let us now denote the conditional, $(p, a) \Rightarrow (p, b)$ by $\mathscr{I}_p(a, b)$. We may say that $\mathscr{I}_p$ is a function $\mathscr{L} \times \mathscr{L} \to \{0, 1\}$ that maps every pair of propositions (a, b) to the conditional $(p, a) \Rightarrow (p, b)$. This transforms the two conditions named above into

$$\mathscr{I}_p(a, b) = 1 \text{ if and only if } a \preccurlyeq b \tag{4}$$

and

$$a \wedge \mathscr{I}_p(a, b) \preccurlyeq b. \qquad \text{Modus Ponens}$$

We have the following theorem:

Theorem 8. *Every orthomodular lattice satisfies* 4 *and* Modus Ponens*, and the following set of conditions for all propositions* a, b *and* c *and for all states* p*:*

1. $a \wedge c \preccurlyeq b$ *implies* $a' \vee (a \wedge c) \preccurlyeq \mathscr{I}_p(a, b)$
2. $a \wedge (a' \vee c) \preccurlyeq b$ *implies* $c \preccurlyeq \mathscr{I}_p(a, b)$ *and* $a \wedge b \preccurlyeq \mathscr{I}_p(a, b)$
3. $a \wedge c \preccurlyeq b$ *implies* $(a \wedge c) \vee (a' \wedge c) \preccurlyeq \mathscr{I}_p(a, b)$

4. $a \wedge c \preccurlyeq b$ *implies* $c \preccurlyeq \mathscr{I}_p(a, b)$ *if* a *and* c *commute.*

Conversely, every lattice that satisfies 4 *and* Modus Ponens *and the set of conditions above is orthomodular, and the conditional function is necessarily* $\mathscr{I}_p$*.*

An orthomodular lattice is an orthocomplemented lattice that satisfies the weak modular law. The theorem shows that orthomodularity corresponds to the 'survival' of the notion of the logical conditional. If we would not have orthomodularity, there would be no function that expresses the conditional for quantum mechanical systems.

8 The EPR Paradox

8.1 Is the EPR Paradox Really a Paradox?

If we recall the explanation of the EPR paradox given in Section 2.5, Einstein, Podolsky and Rosen [6] argued that quantum mechanics is incomplete because there are elements of reality that do not correspond with an element of the theory. We have two entangled particles, and by measuring one we can determine a characteristic - the value of some observable - of the other. This means that there must have been an element of reality in the particles even before measurement. But quantum mechanics was unable to predict this element of reality: we cannot predict the outcome of our measurement. In other words, according to EPR, every element of reality must have a counterpart in physical theory. The theory of quantum mechanics failed to predict some characteristics of the particle that was not measured, but that did lose its superposition and behaved as if it was measured. Thus, they concluded, quantum mechanics must be incomplete.
We might wonder: Has this paradox been solved since EPR published their paper?
EPR tried to solve the paradox by introducing hidden variables that could account for the elements of reality that did not have a counterpart in the theory of quantum mechanics before. However, the physicist John Stewart Bell demonstrated ([2],[4]) that no physical theory with hidden variables could ever reproduce all the predictions made by quantum mechanics.
Although this might seem as something that only complicates the matter further (since now we have to deal with a theory that is incomplete and for which hidden variables offer no solution), Bell showed that no theory with local hidden variables[14] could ever fully predict quantum mechanics. The theory of local hidden variables assumes the principle of locality and the principle of reality to be true. Therefore, one has to give up either locality or realism if one wants to avoid contradictions or incompleteness. Recall that locality is the principle that systems can only interact at short distances. This entails that no influence can propagate faster than the speed of light. With realism we mean the principle that experiments give an outcome independent of the experimenter: Even if the results of the experiment are not known before measurement, as is often the case with quantum mechanics, that does not mean that the act of observation by the experimenter *created* the results. This principle is also known as counterfactual definiteness.
Bell showed that a complete theory of quantum mechanics cannot satisfy both the principle of locality and that of realism.

8.2 Hidden Variables

As we already know, quantum mechanical systems do not always behave in a deterministic way: It is only until one measures a particle that for example its spin will be known. The theory of quantum mechanics cannot predict the spin of a particle before observation[15]. One might ask whether there is some deeper truth, a hidden reality of which we are unaware, that *can* predict the outcomes of quantum mechanical experiments. After all, a probabilistic view of the world can have quite contra-intuitive consequences. As Einstein famously put it, 'I am convinced that God does not play dice.'[16]
Einstein then developed the idea of hidden variables, which would account for the characteristics of the particle that was not measured. This means that some quantity, say λ, is needed to characterize the state of a quantum mechanical system fully. Note here that we do not know what λ looks like (hence the prefix 'hidden'). It might be a number, a set of numbers, a set of functions, anything. Let us assume again that we have two particles that have spin $\frac{1}{2}$, that is, we

[14] Local hidden variables are variables that cannot be influenced instantaneously by events at a great distance (i.e. faster than light).
[15] Quantum mechanics can at most give us a probability distribution of the spin states.
[16] Quote from a letter to Max Born in 1926, see `http://alberteinstein.info/vufind1/Record/EAR000038009`

measure the spin to be in an up-state or in a down-state. This might be a system of an electron and a positron (the singlet state). Bell suggested that instead of measuring the spins along the same direction, one could also measure the spins of the particles along different directions. The vector **a** could denote the unit vector that indicates the direction of the measurement of the electron spin, and **b** then denotes the unit vector in the direction of the measurement on the positron. Each measurement yields either spin up (+1) or spin down (-1)[17]. The probability of obtaining the same result depends on the relative angles at which the spin measurements are performed.
Bell proposed to calculate the average of the product of the measured spins: $P(\mathbf{a}, \mathbf{b})$ would denote this average. If the detectors are antiparallel, the product $P(\mathbf{a}, \mathbf{b})$ would always be +1, since the spins of the electron and positron always point in antiparallel directions. If the detectors are parallel, $P(\mathbf{a}, \mathbf{b})$ is always -1. If the detectors measure in orthogonal directions, $P(\mathbf{a}, \mathbf{b}) = 0$. For *arbitrary* orientations, quantum mechanics predicts

$$P(\mathbf{a}, \mathbf{b}) = -\mathbf{a} \cdot \mathbf{b} = -\cos(\theta), \tag{5}$$

where θ is the angle between **a** and **b**. For a thorough derivation of this equality, see [18]. The above equation expresses the expectation value of the electron being in direction **a** and the positron being in the direction of **b** (cf. Bell, [2]). Bell discovered that this prediction is incompatible with any local hidden variable theory.

8.3 Bell's Inequality

Suppose that influence can only happen locally. The outcome of the measurement of our electron will thus have no influence on the outcome of the positron measurement. Suppose furthermore that the complete state of our system is characterized by the hidden variables λ. Remember, we do not know what λ looks like. We assume that measurement of λ is actually sampling of λ from a (normalized) probability distribution $\rho(\lambda)$. We define $A(\mathbf{a}, \lambda)$ as the function that gives the result of the electron measurement at direction **a**, and $B(\mathbf{b}, \lambda)$ as the function that gives the result of the positron measurement at direction **b**. As we have seen, these functions can only take the values ± 1:

$$A(\mathbf{a}, \lambda) = \pm 1, \quad B(\mathbf{b}, \lambda) = \pm 1. \tag{6}$$

The expectation value of the product of the measurements is

$$P(\mathbf{a}, \mathbf{b}) = \int \rho(\lambda) A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda) \mathrm{d}\lambda. \tag{7}$$

If the detectors are parallel, then

$$A(\mathbf{a}, \lambda) = -B(\mathbf{a}, \lambda),$$

so Equation (7) can be written as

$$P(\mathbf{a}, \mathbf{b}) = -\int \rho(\lambda) A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) \mathrm{d}\lambda.$$

Let **c** be any other unit vector. Then

$$P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c}) = -\int \rho(\lambda) \left[A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) - A(\mathbf{a}, \lambda) A(\mathbf{c}, \lambda) \right] \mathrm{d}\lambda.$$

Since $A(\mathbf{b}, \lambda)^2 = 1$ for all **b**, it follows that

$$P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c}) = -\int \rho(\lambda) \left[1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda) \right] A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) \mathrm{d}\lambda.$$

[17] Here we work in units of $\hbar/2$, as we have done before

And because of Equation (6), we have that $-1 \leq A(\mathbf{a}, \lambda)A(\mathbf{b}, \lambda) \leq 1$, so that

$$|P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})| \leq \int \rho(\lambda) \left[1 - A(\mathbf{b}, \lambda)A(\mathbf{c}, \lambda)\right] \mathrm{d}\lambda.$$

But the second term on the right hand side is just $P(\mathbf{b}, \mathbf{c})$, so that

$$1 + P(\mathbf{b}, \mathbf{c}) \geq |P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})|. \qquad \text{Bell inequality}$$

This inequality holds for all hidden variable theories; we have no restrictions on λ or ρ. Armed with this inequality, we may finally construct a system in which we have a contradiction with (5).
Suppose all three vectors lie in the plane, and $\mathbf{c}$ makes a $45°$ angle with $\mathbf{a}$ and $\mathbf{b}$. From Equation (5) it follows that

$$P(\mathbf{a}, \mathbf{b}) = 0 \quad \text{and} \quad P(\mathbf{a}, \mathbf{c}) = P(\mathbf{b}, \mathbf{c}) = -\cos(\pi/4) \approx -0.707.$$

But from Bell's inequality we have

$$0.707 \not\leq 1 - 0.707 = 0.293.$$

And so we reach our conclusion: Any hidden variable theory is incompatible with quantum mechanics.

So, as it turns out, we must give up either realism (the results are independent of the observer) or locality (the premise that the outcome of a measurement on the electron did not influence the outcome of a measurement on a positron) in order to create no contradictions. The refutation of locality meant that systems *could* influence each other at large distances. This seems odd at first. Einstein said no information can travel faster than the speed of light, but in the EPR experiment one particle was sent to the other side of the universe and its state was known instantly after measurement. However, this leads to no contradiction. The crux is the quantum mechanical property that we cannot predict the outcome of a measurement beforehand. That is, we cannot *steer* the outcome in any way. If we try to influence the states of the particles, for example by means of electromagnetic waves at a specific frequency or by using polarizers on photons, we actually still measure them. Then all states are specified and the entanglement is lost. There is no causal influence that travels faster than light.

9 Concluding Remarks

In this Bachelor thesis we have tried to understand and explain the mathematical principles of quantum mechanics. We have looked especially at the famous paradox by Einstein, Podolsky and Rosen that caused physicists to drastically revise their way of viewing the world.

First we have acquainted several concepts from the field of quantum mechanics such as superposition, locality and entanglement, and then we have encountered the notions of states, observables and propositions. We have seen that these propositions form a complete, orthomodular lattice, noted as $\mathscr{L}$. Through a construction that uses projective geometries and the reducibility of lattices, we can associate the lattice of propositions to a lattice of subspaces of a Hilbert space. Propositions then turn from yes-no-experiments to subspaces of a Hilbert space. Observables, which were seen as σ-homomorphisms in the lattice of propositions, become spectral measures (or, equivalently, self-adjoint linear operators) and states turn from measures on $\mathscr{L}$ to functionals on the projection operators.
After this, we explored the formal quantum logic as proposed by Birkhoff and Neumann (in their famous paper from 1936, see [20]). This gave us some insight in the connection between orthomodularity and the logical conditional. Finally, we arrived at the Einstein-Podolsky-Rosen paradox. We tried to explain why this paradox has fallacies, and how John Stewart Bell proved that the paradox cannot be saved by introducing hidden variables. As it turns out, we have to give up realism or locality to make quantum mechanics a complete and sound theory.

There is still much debate in the field of quantum mechanics. There are scientists that believe that we should give up the principle of reality: When we measure a quantum mechanical system, the observer *makes* the system choose one of its possible states, which is then the measured state. Others say the principle of locality cannot be true: Systems may interact over great distances (but they never contradict Einstein's law that nothing travels faster than the speed of light). Some even suggest that we give up both. The field of quantum mechanics is still relatively young, and its complex nature makes it very hard to comprehend completely. Of course, the fact that most of quantum mechanics is situated at nanoscale means that it is quite difficult to get a good view of what happens exactly. So it is no wonder that physicists are still arguing about the nature of particles and about the interactions they have.
The subject of quantum logic is not undisputed, either. In fact, it is difficult to find two articles that agree with each other. The quantum logic of Birkhoff and Von Neumann is heavily criticized for its need to conserve the law of the excluded middle. Proponents of logical intuitionism as founded by Brouwer and Heyting might reject this law. They say that, in the same way that Schrödingers cat is dead and alive at the same time, some propositions should be able to be true and not true simultaneously as well.

As in all sciences, quantum mechanics is a field where certainties are rare, but it is this uncertainty that makes quantum mechanics so fascinating to discover. It has many faces and as we have seen throughout this thesis, many types of mathematics are used to describe it: Algebra, logic, linear algebra, analysis, probability theory and statistics, functional analysis and so on. For me, this was the proof that quantum mechanics is absolutely worth exploring.

9.1 Acknowledgements

This Bachelor thesis could not have been written without the help and encouragement from others. First of all, I would like to thank dr. K. P. Hart and dr. Eva Coplakova for their time and patience, helping me with the difficult parts of logic and mathematics and giving me direction where needed. Prof. dr. Jan van Neerven also really encouraged me with his great enthusiasm and helped me with the more physical (and sometimes philosophical) side of

mathematics. I also owe credit to dr. ir. Silvania Pereira of the faculty of TNW for answering some questions about quantum mechanics.
Finally, I would like to thank my family and friends, for being there for me and giving great advice. Especially Pim, Arthur, Rik and Norma for proofreading my script and Gijsbert and Joris, for advice and support.

References

[1] Reinhold Baer. *Linear Algebra and Projective Geometry*. Academic Press, 1952.

[2] John S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195–200, 1964.

[3] Enrico G. Beltrametti and Gianni Cassinelli. *The Logic of Quantum Mechanics, Encyclopedia of Mathematics and its Applications*, volume 15. Addison-Wesley Publishing Company, 1981.

[4] Guy Blaylock. The EPR Paradox, Bell's Inequality and the Question of Locality. *American Journal of Physics*, 78(1):111–120, 2010.

[5] Paul Dirac. *Principles of Quantum Mechanics*. Cambridge University Press, 1958.

[6] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review*, 47:777–780, 1935.

[7] Euclid. Elements. `http://aleph0.clarku.edu/~djoyce/java/elements/bookI/bookI.html`.

[8] I. M. Gelfand and M.A. Naimark. On the Imbedding of Normed Rings into the Ring of Operators in Hilbert Space. *Receuil Mathématique*, 12(54):197–213, 1943.

[9] Andrew M. Gleason. Measures on the Closed Subspaces of a Hilbert space. *Journal of Mathematics and Mechanics*, 6(6):885– 893, 1957.

[10] David J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Education, second edition, 2005.

[11] R.I.G. Hughes. *The Structure and Interpretation of Quantum Mechanics*. Harvard University Press, 1989.

[12] Josef M. Jauch. *Foundations of Quantum Mechanics*. Addison-Wesley Publishing Company, 1968.

[13] Richard V. Kadison. Notes on the Gelfand-Neumark Theorem. *Contemporary Mathematics*, 167:21–29, 1994.

[14] Sifis Petrakis. Introduction to Banach Algebras and the Gelfand-Naimark Theorem. `www.math.lmu.de/~petrakis/INTRODUCTION%20TO%20BANACH%20ALGEBRAS.pdf`.

[15] Constantin Piron. Axiomatique Quantique. *Helvetica Physica Acta*, 37(4-5):439, 1964. `http://tinyurl.com/loxfpg5`.

[16] Constantin Piron. *Foundations of Quantum Physics*. W. A. Benjamin, Inc., 1976.

[17] H. Putnam. Is Logic Empirical? *Mathematics, Matter and Method*, pages 174–197, 1975.

[18] Kevork Spartalian. About spins. `www.uvm.edu/~kspartal/EPR_Simulation/AboutSpins.pdf`.

[19] V.S. Varadarajan. *Geometry of Quantum Theory*. Springer, 2007.

[20] John von Neumann and Garrett Birkhoff. The Logic of Quantum Mechanics. *Annals of Mathematics*, 37(4):823–843, 1936.

Index