Quantum-mechanical wave functions as elements of reality

Esger Renkema

19 June 2013

A bachelor thesis
written under supervision of

Prof. dr. D.G.B.J. Dieks

Institute for History and Foundations of Science
Departement of Physics, Universiteit Utrecht

Abstract

This thesis explores the relations between the wave functions that describe the state of a system in quantum mechanics and the hypothetical »ontic« states a system can have in reality. What quantum mechanics precisely implies about the nature of reality has always been a disputed point; among the contended features of quantum mechanics as a physical theory are completeness and locality. Recent results indicate that if some assumptions are made about reality being local, then quantum-mechanical and ontic states must be in one-to-one correspondence.

Contents

1 Introduction

To someone studying physics, it often seems to be a natural way of thinking that physical theories are not merely tools that are used to predict outcomes of measurements, but representations of the fundamental nature of the universe or »reality«. Conversely, the universe is often supposed to be of the same mathematical structure as the theory of physics that describes the outcomes of experiments within it and the mathematical objects of the theory are colloquially said to correspond to objects »in existence«. In the mind of the physicist, furthermore, the measurable quantities that are the proper subjects of physical theories are often seen as *properties* of the measured system and these properties are thought to be contained in or determined by the system itself.

And so, the gravitational force is thought to be an object in the »real world« and not merely a mathematical trick for determining the time at which a thrown apple hits the head of a sleeping colleague – a thrown apple that, according to intuïtion, should have the properties »place« and »velocity« in addition to »ripeness« and »being an apple«.

One of the problems that physicists encountered at the beginning of the previous century was that it seemed impossible to reconcile the upcoming theory of quantum mechanics with the classical intuïtion mentioned above, for the mathematical structure of quantum mechanics itself seems to run contrary to the idea of systems having well-defined properties the way they have in classical mechanics. But although this »realistic« understanding of a physical theory is in no way required by physics itself – it simply is a convenient and intuïtive way of thinking about physics – it also seems unreasonable to put away these ideas as useless or »naïve«, since physics is, after all, the study of the »real world«.

This discrepancy between quantum mechanics and classical intuïtion raised several questions, most famously whether quantum mechanics was a *complete* theory, but also – assuming the completeness of quantum mechanics – what the actual and fundamental properties *were* of a given system. One well-known supporter of classical intuïtion was Albert Einstein and we will come to one of his arguments shortly. First, however, a short description of quantum mechanics might be in order.

1.1 Determinism

In quantum mechanics, the state of every system is described by a complex-valued *wave function*. These wave functions are *probability amplitudes* for the outcomes of measurements performed on the system. The wave functions of quantum mechanics are much like waves on the surface of a pond in that their evolution has a similar mathematical form; and though the wave functions of quantum mechanics can be rather more complex, the same intuïtion applies to both.

According to the standard (or *orthodox*) interpretation of quantum mechanics, the measurement of some physical quantity usually changes the state of the measured system in a process that is known as the *collapse of the wave function.* In mathematical terms, this »collapse« is a sudden change of the wave function to an eigenstate of the (hermitian) operator that corresponds to the measurement. The chance of collapsing to a specific eigenstate can be calculated from the wave function: it is the squared modulus of the inner product between the wave function and the eigenstate. In formulas, that is $|\langle\psi|\alpha\rangle|^2$ for wave function ψ and eigenstate α.

These eigenstates need not form a continuous spectrum; the spectrum might be discrete. This is the »quantum« part of quantum mechanics: some physical quantities can only have values that are integer multiples of a smallest possible value, the *quantum* of that quantity. If the wave function of a system is an eigenstate of the measurement operator, then the outcome of the measurement is certain, since the inner product of the wave function with itself equals 1. In general, however, the wave function is not an eigenstate of the to be quantity that is to be measured. In that case, the the result of the measurement is uncertain, with the chances for each possible result given by the inner product mentioned above. Because of this, quantum mechanics is not a *deterministic* theory.

Related to this indeterminism of quantum mechanics is the phenomenon of *complementary* physical quantities. Because two measurement operators might have different eigenstates and because the measurement of one property changes the wave function to an eigenstate of the measurement operator, there exist physical quantities that cannot both be known at the same time – after the first measurement is performed, the system is changed to a state that can correspond to several possible outcomes of the second measurement. In mathematical terms, two measurements are complementary if their operators do not commute. Examples of complementary physical quantities are position and momentum, or two different spin directions.

1.2 Completeness

As stated above, quantum mechanics can be considered indeterministic because outcomes of measurements are not uniquely determined by the quantum-mechanical state of the measured system. Uncertainty *does* occur in classical mechanics, but there it is always a result of the observer having a shortage of information. Take for example the erratic motion of small particles suspended in a liquid: in classical statistical physics, this

Brownian motion is random because only the temperature and pressure of the fluid are known, intensive properties that are determined by averaging over all particles in the fluid. If you could account for the movements of all particles, the motion would become deterministic. (And of course, classical thermodynamics assumes that every particle *has* an exact place and velocity. This is not the case in quantum mechanics.)

Therefore, comparing the old and the new mechanics it is a natural question to ask whether the indeterminism of quantum mechanics is fundamental and unavoidable, or whether quantum mechanics is somehow *incomplete*, meaning that there exist as of yet hidden variables that determine the system more completely and that give better predictions to the outcomes of measurements. If that were the case, quantum uncertainty – or at least part of it – would not be a fundamental property of the universe, but an expression of a lack of knowledge in the mind of the experimenter.

1.3 Realism

When discussing the completeness of quantum mechanics, it is often inescapable to assume the existence of a »real world«. The idea that objects in a physical theory can correspond to objects in reality is called *realism*. Realism implies that a system has physical quanties of its own, that is: physical quantities that exist also if the system is unobserved.

In realistic terminology, the »real« states of a system are called *ontic* states, after the present participle of *ἐιμί* (»to be«). Surprisingly, it will turn out to be possible to reason about these ontic states by assuming the correctness of the predictions made by quantum mechanics. The second part of this thesis will analyse some recent publications on this matter.

Regarding complementary observables, an old question is whether two complementary observables can have *simultaneous reality*; this is often understood as both quantities having a definite value. It was the subject of the famous paradox of Einstein, Podolsky and Rosen,[1] who tried to prove the incompleteness of quantum mechanics by connecting complementary quantities with different ontic states. Ultimately, their attempt was unsuccesful; for in 1964, Bell[3] showed that reasoning about simultaneous values of complementary observables led to a contradiction with quantum mechanics that could be decided experimentally. The first part of this thesis will consist of a treatment of this discussion.

2 The EPR paradox

In 1935, Einstein, Podolsky and Rosen published the critique of quantum mechanics[1] that became known as the *EPR paradox*. The central question to their argument was *can two noncommuting observables have simultaneous reality?* meaning to ask whether it is possible to think of both observables as having a definite values. Since quantum mechanics contains no notion of a system being in differing eigenstates of two noncommuting observables, they reasoned that if that were the case, then quantum mechanics must be incomplete.

Underlying the paradox is the assumption of *local realism*. Realism has been treated above, we will come to a definition of locality shortly.

As a starting point for their argument, Einstein, Podolsky and Rosen take the idea of an *element of reality*, that is: some aspect of the ontic state. It is not possible to reason about elements of reality directly, since the incompleteness of quantum mechanics will have to be proved by showing that there exist elements of reality not accounted for; therefore, the elements of reality have to be assumed (yet) unknown. However, Einstein, Podolsky and Rosen argue that if it would be possible to exactly predict the value of a physical quantity, then that at least would indicate the existence of a corresponding element of reality.

Now according to quantum mechanics, it is sometimes possible to know the outcome of a measurement in advance: this is the case if the measured system already is in an eigenstate of the measurement operator. Thus, a single observable can, in some cases, correspond to an element of reality. For two noncommuting observables, this need not be the case: while it is possible to know the outcome of the second measurement with certainty too, this can only be done by changing the system and this change might destroy the connection between the first observable and reality.

EPR propose to work around this problem with the phenonenon of *entangled states*. Entangled states occur because in quantum mechanics, the combined state of two particles is not restricted to the products of two single-particle states, but to any linear combination of them. This makes it possible to construct combined states where the outcomes of measurements on the individual particles are strongly correlated. For instance, it is possible to construct a system of two photons where if the polarisation of one photon is measured the outcome is random, but where the polarisation measurements of both photons always agree. (Polarisation can be measured in different directions and these measurements do not commute. We will use that fact shortly.) If it is possible, like in this example, to exactly know the outcome of a measurement on the second particle by measuring the first, the two particles are said to be maximally entangled.

Physically speaking, an entangled state can occur when two particles interact and then get separated. This separation can get arbitrarily large without altering the entanglement. The large separation that is possible between entangled particles plays an important role when it comes to the *locality* of physics. Briefly, the assumption that physics is local means that no physical process propagates faster than the velocity of light. This means that for two entangled particles that are far-apart, measuring one particle (and thus changing its wave function) must leave intact the ontic state of the other particle. The meaning of locality will be explored further in section 2.2.

Assume locality and consider a system of two entangled photons. Because the photons are entangled, all polarisation measurements made on both particles must be in agreement. Thus, if you would measure the horizontal polarisation on the first particle, you would know for certain the outcome of the same measurement on the second particle. This means that – after a measurement on the first particle – the horizontal polarisation of the second particle must correspond to an element of reality. Because of locality, however, performing a measurement on the first particle cannot change the (ontic) state of the second particle. This means that the horizontal polarisation of the second particle must also correspond to an element of reality if the measurement on the first particle is not (yet) performed.

Because the only thing that was needed to establish the correspondence of horizontal polarisation of the second particle with an element of reality was the *possibility* of measuring the horizontal polarisation of the first particle, the argument can be repeated for polarisation in the vertical direction to show that vertical polarisation should correspond to an element of reality, too. The two quantities, therefore, must have simultaneous reality.

Now in quantum mechanics, measurements of horizontal and vertical polarisation do not commute, so that it is impossible to represent both quantities at the same time. Einstein, Podolsky and Rosen conclude from this that quantum mechanics is incomplete.

2.1 Joint and marginal distributions

From the premise of local reality, the EPR paradox concludes that it is possible to reason about simultaneous values of noncommuting observables. This will turn out to be of crucial importance, since from this proposition, the Bell inequalities can be derived. Before doing that, it is necessary to develop some mathematical notation. That will be done in the form of *joint and marginal probability distributions*. This notation will be used throughout the rest of this thesis.

Consider measuring the spin of a simple electron. Quantum mechanics dictates that, since measurements of two different spin directions do not commute, the outcomes of two different measurements on the same electron cannot both be known in advance. Quantum mechanics does, however, give you a probability distribution for solitary measurements, describing their outcomes. Let us call these outcomes σ_i with $\sigma \in \{\uparrow, \downarrow\}$ and i labelling the measurements, so we can write $\mathrm{P}_i(\sigma_i)$ for the distribution of outcomes of the ith measurement.

Given two such probability distributions, you might try to contruct a joint probability distribution $\mathrm{P}_{12}(\sigma_1, \sigma_2)$ from which the single measurement distributions can be derived as marginals, that is: for which

$$\sum_{\sigma_2} \mathrm{P}_{12}(\sigma_1, \sigma_2) = \mathrm{P}_1(\sigma_1)\,, \tag{1}$$

holds. Another way to look at this equation is: »P_{12} reduces to P_1 if you know nothing about the outcome σ_2.«

The existence of joint probability distributions is deeply linked to the question whether complementary physical quantities can have simultaneous reality. The existence of a joint distribution for two complementary quantities suggests by itself that this is the case: $\mathrm{P}_{12}(\uparrow_1, \uparrow_2)$ would be the chance of both spins being 'up'. And vice-versa, if you assume that complementary quantities *have* simultaneous values, a joint probability has to exist between them.

It should be noted that complementary quantities having *definite* values is not a necessary condition for a joint distribution to exist; the existence of a joint probability is also implied if the values of the individual quantities are given by chance. The only necessary assumption is that is is possible to reason about the values of both quantities at the same time.

It should also be noted that joint probability distributions of *commuting* observables are a normal and well-defined part of quantum mechanics, since quantum mechanics has no problems with assigning simultaneous values to non-complementary physical quantities. In particular, if measurements 1 and 2 are measurements on different particles, $\mathrm{P}_{12}(\sigma_1, \sigma_2)$ can be calculated directly in quantum mechanics. This joint distribution $\mathrm{P}_{12}(\sigma_1, \sigma_2)$ is not, in general, equal to $\mathrm{P}_1(\sigma_1) \cdot \mathrm{P}_2(\sigma_2)$ since, for instance, the two particles might be entangled.

Garg and Mermin[6] have described the EPR paradox as being a way of determining a joint probability of two noncommuting observables. Because of the conclusion of the paradox that two complementary physical quantities have simultaneous reality, the result of a measurement performed on one particle in a maximally entangled pair must also determine the value of the same physical quantity on the other particle, without disturbing its state. Therefore, the joint probability can be determined by measuring one quantity on the first particle and the other on the second, giving something like a simultaneous measurement of two complementary quantities on one particle.

2.2 Locality

As briefly stated in section 2, locality holds that no physical process can propagate faster than the velocity of light. The assumption of locality is an important premise of the EPR paradox: without locality, it could be possible that performing a measurement on one particle in an entangled pair determines the nature of reality for the second particle. This would invalidate the paradox.

Locality is a common assumption and it will not be challenged in the rest of this thesis, but we have to explore its meaning, since the assumption can be made in different ways.

The first thing to note is that it is not locality that is needed, but a weaker condition of *non-disturbance*: the only real requirement is that measuring one particle does not disturb the state of another particle that might be entangled with the first. Of course, non-disturbance follows immediately from locality.

It is non-trivial to reason about non-disturbance within quantum mechanics, since combined states cannot in general be written as products of single-system states. Instead, combined states can be linear combinations of products of single-system states, so that the idea of changing just the state of one of two entangled particles is a bit problematic. The solution is to use a *partial trace* to obtain a workable idea of the state of a single particle. Using a partial trace, the state of particle $\mathcal{A}$ in the combined state $S \in \mathcal{A} \otimes \mathcal{B}$ can be defined as

$$S^{\mathcal{A}} := \mathrm{Tr}^{\mathcal{B}}(S) = \sum_i \langle \beta_i | S | \beta_i \rangle \qquad (2)$$

with $\{\beta_i\}$ a basis on $\mathcal{B}$. Using this notion of the state of one part of a combined system, it is not very difficult to see that non-disturbance holds in quantum mechanics: if a measurement on particle $\mathcal{B}$ makes its wave function collapse while the result of that measurement remains unknown to the observer of particle $\mathcal{A}$, then the best predictions the observer at $\mathcal{A}$ can make will remain the predictions he makes based on the partial trace. A thorough treatment of the partial trace can be found in [8].

The assumption of non-disturbance is often made together with the assumption that it is possible to reason about simultaneous values of noncommuting physical quantities. In a way, however, the second assumption implies the first. The simultaneous reality of the complementary quantities can be expressed by a joint probability distribution; this joint distribution must have the quantum-mechanical probabilities as marginals, according to definition (1). In that equation, you can see that the probability $\mathrm{P}_1(\sigma_1)$ remains the same if σ_2 is either unmeasured or measured but unknown, in much the same way as non-disturbance in quantum mechanics follows from the definition of the single-particle state that can be found with a partial trace.

Strictly speaking, the above does not prove that reality is local, since there might exist elements of the ontic state that *do* change instantly at the measurement of a far-away entangled particle. These non-local elements of reality, however, cannot influence the outcome of measurements, so that there must exist a model of reality that is local and that fits all experimental results.

3 The Bell inequalities

The publication of the EPR paradox did not quite settle the question whether quantum mechanics was or was not incomplete: the completeness of quantum mechanics could be rescued by challenging one of the assumptions that underlie the derivation of the paradox. This could be either the idea that predicting measurement outcomes with certainty would point to an element of reality – this seems to be the view held by Bohr[8] – or the locality of physics. Allowing quantum mechanics to be non-local is not as problematic as it might seem, since even if quantum states might react instantaneously to far-away events, superluminal transmission of information remains impossible because of the the non-disturbance explained in section 2.2.

Despite the rather fundamental questions it raised, for a long time the EPR paradox seemed to be not very important for the functioning of quantum mechanics. After all, the EPR paradox was not constructed in an attempt to show that quantum mechanics was *wrong*, but in an attempt to show that quantum mechanics was *incomplete*. And although the main result of the arguments behind the paradox – that two complementary physical quantities can have simultaneous reality – led to the conclusion that quantum mechanics was incomplete, this result appeared to be not experimentally testible.

It was only in 1964 that Bell[3] realised that this simultaneous reality of complementary quantities *did* have actual consequences for predictions made about the outcomes of certain measurements. Using the same assumptions made by Einstein, Podolsky and Rosen, he derived an inequality that should hold in any theory making these assumptions, but that was formulated in terms that could be calculated within ordinary quantum mechanics. Subsequently, he showed the existence of a configuration of particles and measurements that made the calculations of quantum mechanics contradict the general inequality.

It has been shown by Fine[5] that the only assumption needed to derive the Bell inequalities is the existence of a joint probability of two noncommuting observables. As mentioned in section 2.1, simultaneous reality of the physical quantities represented by the observables implies the existence of such a joint probability distribution.

In the language of joint probability distributions, the incompatibility between quantum mechanics and the conclusions of the EPR paradox stems from the mathematical fact that given an arbitrary set of multivariable probability distributions, it is not always possible to find a joint probability that gives the smaller probability distributions as marginals. In more concrete terms this means that the joint probabilities of commuting quantities that occur in quantum mechanics *do not allow the mathematical existence of a joint probability of all involved quantities.* Thus, no model of reality can assign simultaneous probabilities to noncommuting observables without contradicting quantum mechanics.

Garg and Mermin[6] have thoroughly explored the mathematics behind joint probability distributions and give a general method of determining whether a joint probability distribution can exist with a given set of marginals. The Bell inequalities, however, are a simpler way of proving that the probabilities given by quantum mechanics not always allow a joint probability. In the next section, we will derive a Bell inequality that was first formulated by Clauser, Horne, Shimony and Holt[4]. The derivation itself follows Braunstein and Caves[7] and starts with only a joint probability distribution.

3.1 Deriving the Bell inequalities

Take two electrons as object to reason about, and choose two (different) spin measurements on each particle. Say that σ_1 and σ_3 are the outcomes of the measurements on the first and σ_2 and σ_4 the outcomes for the second electron. Now in quantum mechanics, it is not meaningful to use σ_1 and σ_3 (or σ_2 and σ_4) in the same expression, since they are the results of complementary measurements. Only $\mathrm{P}_{12}(\sigma_1, \sigma_3)$, $\mathrm{P}_{23}(\sigma_1, \sigma_4)$, $\mathrm{P}_{34}(\sigma_2, \sigma_3)$ and $\mathrm{P}_{41}(\sigma_2, \sigma_4)$ exist as quantum-mechanical joint distributions, since measurements performed on different particles *do* commute.

Assume the existence of a grand joint probability distribution for all four measurement results, $\mathrm{P}_{1234}(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$, from which the above distributions can be derived:

$$\mathrm{P}_{12}(\sigma_1, \sigma_2) = \sum_{\sigma_3, \sigma_4} \mathrm{P}_{1234}(\sigma_1, \sigma_2, \sigma_3, \sigma_4) \tag{3}$$

and so on for the other three well-defined joint distributions.

Since the goal of this derivation should be an expression that can be compared with the predictions of quantum mechanics, our final statement may only depend on the four two-variable distributions mentioned above. Define for $a \in \{1, 3\}$, $b \in \{2, 4\}$ a correlation function

$$\mathrm{C}_{ab} = \sum_{\sigma_a, \sigma_b} \sigma_a\, \sigma_b\, \mathrm{P}_{ab}(\sigma_a, \sigma_b) \tag{4}$$

to contstruct the following formula

$$|\mathrm{C}_{12} + \mathrm{C}_{23} + \mathrm{C}_{34} - \mathrm{C}_{41}| \tag{5}$$

that can be computed using ordinary quantum mechanics. Note that we need to assign a numerical value to the measurement outcomes σ_i. As usual, let the 'up' outcome be $+1$ and the 'down' outcome be -1.

Using the grand joint probability (3), it is possible to establish an upper bound on expression (4) that, in the end, will conflict with quantum mechanics. To do that, note that if you can use P_{1234}, the definition of a marginal (3) can be substituted into the definition of C_{ab} (4) to give a new form that does not depend anymore on the quantum-mechanical joint probabilities:

$$\mathrm{C}_{ab} = \sum_{\sigma_1, \sigma_2, \sigma_3, \sigma_4} \sigma_a\, \sigma_b\, \mathrm{P}_{1234}(\sigma_1, \sigma_2, \sigma_3, \sigma_4)\ . \tag{6}$$

Using this equation, expression (5) can be rewritten into the following formula:

$$|\sum_{\sigma_1, \sigma_2, \sigma_3, \sigma_4} (\sigma_1(\sigma_2 - \sigma_4) + \sigma_3(\sigma_2 + \sigma_4))\ \mathrm{P}_{1234}(\sigma_1, \sigma_2, \sigma_3, \sigma_4)\,|\ . \tag{7}$$

This expression might look a bit daunting at first, but it can be simplified greatly by eliminating the measurement outcomes σ_i: because the values of these outcomes can only be ± 1, the part of (7) that contains the outcomes can only take two values:

$$(\sigma_1(\sigma_2 - \sigma_4) + \sigma_3(\sigma_2 + \sigma_4)) = \pm 2\ . \tag{8}$$

Then, the absolute value and the summation can be interchanged to give the following inequality:

$$[\text{expression (5)}] \leq \sum_{\sigma_1,\sigma_2,\sigma_3,\sigma_4} |\pm 2| \ \mathrm{P}_{1234}(\sigma_1,\sigma_2,\sigma_3,\sigma_4) \ . \tag{9}$$

Adding up all values of P_{1234} wil just give 1, so what we have here is a fixed upper bound on (5):

$$|\mathrm{C}_{12} + \mathrm{C}_{23} + \mathrm{C}_{34} - \mathrm{C}_{41}| \leq 2 \ , \tag{10}$$

which is the sought-after Bell inequality.

3.2 Violating the Bell inequalities

Although the Bell inequality (10) was derived using a grand joint probability distribution, its final form only depends on correlation functions C_{ab} that can be calculated in quantum mechanics (that is: for which a and b are not both even or odd). All that is left to do now, is to show that for some experimental setups, this equation does not hold when calculated using the quantum-mechanical joint probabilities.

Such an experimental setup can easily be found for two maximally entangled electrons. A combined state of two maximally entangled electrons has the quantum state $(\frac{1}{2}\sqrt{2})(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$: the result of a measurement on one electron must be the opposite of the result of the same measurement on the other electron.

Remember that measurement numbers 1 and 3 are spin measurements on one particle and that numbers 2 and 4 are spin measurements on the other particle. For simplicity, let the measurement directions lie in a plane with angle θ between subsequent measurement numbers. For such a setup, C_{12}, C_{23} and C_{34} are equal. Their value can be calculated from the corresponding quantum-mechanical joint distributions:

$$2\ \mathrm{P}_{12}(\sigma_1,\sigma_2) = \begin{cases} \sin^2\dfrac{\theta}{2} & \text{for } \sigma_1 = \sigma_2 \\ \cos^2\dfrac{\theta}{2} & \text{for } \sigma_1 \neq \sigma_2 \end{cases} \tag{11}$$

Note that this function contains $\theta/2$ instead of θ: the factor of $\frac{1}{2}$ arises from the fact that while the spatial directions 'up' and 'down' differ by an angle of π, their corresponding basis vectors $|\uparrow\rangle$ and $|\downarrow\rangle$ lie $\pi/2$ radians apart in Hilbert space. By putting (11) in equation (4), we arrive at the value of the first three correlation functions:

$$\mathrm{C}_{12} = -\cos\theta \tag{12}$$

Of course, the last correlation function is given by $\mathrm{C}_{41} = -\cos(3\theta)$, calculated in the same way.

With the four correlation functions calculated, the left side of Bell inequality (10) equals

$$|3\cos(\theta) - \cos(3\theta)|\,. \tag{13}$$

It is easily verified that this expression reaches its maximum value of $2\sqrt{2}$ at $\theta = \pi/4$, which clearly exceeds the maximum value of 2 given by the Bell inequality.

3.3 Chained Bell inequalities

The Bell inequality presented above is sufficient for showing the incompatibility between quantum mechanics and the realistic view of Einstein, Podolsky and Rosen. Laying to rest the EPR paradox, this section will focus on exploring the meaning of the Bell inequalities within quantum mechanics.

The preceding sections reason about an inequality that contains four measurements. It is not difficult to construct similar inequalities that consider more than four; one method of expanding the Bell inequality is by *chaining* it. Braunstein and Caves[7] found, perhaps surprisingly, that the process of chaining Bell equations produces *stronger* conditions than can be reached using only four measurements.

To construct a chained Bell inequality, take $2N$ measurements instead of the four measurements of equation (10). Have the odd numbered measurements apply to one particle and the even numbered measurements to another. Using the same notation as above, the following $N/2 - 1$ Bell inequalities

$$\left|\mathrm{C}_{1i} + \mathrm{C}_{i(i+1)} + \mathrm{C}_{(i+1)(i+2)} - \mathrm{C}_{(i+2)1}\right| \leq 2 \tag{14}$$

can be added up to give the chained Bell inequality

$$\left|\mathrm{C}_{12} + \mathrm{C}_{23} + \ldots + \mathrm{C}_{(2N-1)(2N)} - \mathrm{C}_{(2N)1}\right| \leq 2N - 2 \tag{15}$$

that still incorporates all measurements in a way that is compatible with quantum mechanics.

Constructing a system that violates this chained Bell inequality can be done in the same way as for a normal Bell inequality. If θ is again the angle between neighbouring measurements, the expression for the left side of (15) becomes

$$|(2N-1)\cos\theta - \cos(\theta(2N-1))|. \tag{16}$$

This expression has its maximum value of $2N\cos(\pi/2N)$ at $\theta = \pi/2N$. As N nears infinity (N was fixed at the beginning of this derivation, but could have been arbitrarily high), the value of (16) according to quantum mechanics nears $2N$. This is a stronger violation than the $2\sqrt{2}$ that could be reached for (10).

The Bell inequalities were derived to show that no theory compatible with quantum mechanics may assign joint probabilities to complementary physical quantities, thereby putting severe constraints on the construction of hidden-variable theories. Chaining the Bell inequalities does not add to that proof, but it does help in understanding the property of quantum mechanics that makes these joint probabilities impossible.

To see this, explore the experimental setup maximizing (16) that we derived in a rather sterile, mathematical way. It consists of a series of measurement operators, with two neighbouring operators operating on different particles. The angle of measurement between subsequent operators is very small, thus maximizing the first $2N-1$ correlation functions in (15). The last correlation function is maximized (and made positive) by the *total* angle between measurements.

What this illustrates, is that in quantum mechanics, correlations are very strong at small angles. These strong correlations arise because the probability amplitudes have, for small angles, a linear dependence on θ, thus making the dependence of the actual probabilities quadratic. Braunstein and Caves draw no further conclusions from this.

4 Hidden variables

In the preceding sections the focus has, ultimately, always been with the EPR paradox. But although the Bell inequalities prove that the assumptions that underlie that paradox contradict quantum mechanics, the actual question that the EPR paradox sought to address – whether quantum mechanics was complete – cannot be answered by the Bell inequalities.

The question whether quantum mechanics is complete can be studied by examining *hidden-variable theories*, theories extending quantum mechanics with a set of variables that help determine the measurement results but that have been hidden until now.

One example of such a hidden variable theory has been constructed by Bohm[2]. The Bell inequalities do not apply to Bohm's theory, but although the theory is realistic, it is not local. As stated in 2.2, in this thesis we will only consider local theories.

In a recent publication, Colbeck and Renner[11] have used a form of the chained Bell inequalities to argue that quantum mechanics is »complete« in the sense that no alternative theory can make both compatible and more accurate predictions about the outcomes of measurements. Their argument does not require the existence of a joint probability distribution, but it needs the locality that is outlined in section 2.2. We will spend some time duplicating parts of Colbeck and Renner's results using the language and notation introduced above.

4.1 Physical theories

Before we start reasoning about alternative physical theories, it might be prudent to define the precise meaning of »physical theory«. In the argument of Colbeck and Renner, a physical theory can be seen as a function that maps the combination of some notion of *state* and the specification of a *measurement* to a probability distribution over the space of possible outcomes of the measurement.

We will label the states of quantum mechanics – wave functions – with ψ and the states of our rivalling theory with λ. If we want to be able to relate our new theory to quantum mechanics, we need to assume the existence of a joint probability $\mathrm{R}(\lambda|\psi)$ between the states of the two theories. This assumption is implicit in the arguments of Colbeck and Renner.

The existence of the conditional distribution $\mathrm{R}(\lambda|\psi)$ implies that ψ and λ have simultaneous reality. This is reasonable to assume, since otherwise the predictions made by both theories cannot be compared. Using $\mathrm{R}(\lambda|\psi)$, we can formulate a requirement for the two theories being compatible: if $\mathrm{P}_i^{\psi}(\sigma_i)$ is (for some experiment labeled i) the distribution of outcomes predicted by quantum mechanics and $\mathrm{P}_i^{\lambda}(\sigma_i)$ the distribution predicted by the new theory,

$$\mathrm{P}_i^{\psi}(\sigma_i) = \sum_{\lambda} \mathrm{P}_i^{\lambda}(\sigma_i)\ \mathrm{R}(\lambda|\psi) \tag{17}$$

has to hold. It can be stated intuitively as »when you know ψ but do not know λ, quantum mechanics arises from averaging over the other theory«.

The final assumption made by Colbeck and Renner is that the hidden state λ may not depend on the choice of what to measure. This condition of non-disturbance follows directly from locality; it also holds in quantum mechanics. It is a very reasonable assumption: without it, it would not be possible to compare, for a system in a given state λ, the predictions made about two different measurements $\mathrm{P}_a^{\lambda}(\lambda)$ and $\mathrm{P}_b^{\lambda}(\lambda)$, because the choice of a and b would have changed the state λ, thereby turning the two probability distributions into statements about differing systems.

4.2 Variational distance

Given two physical theories, how can we say anything about the theories being (essentially) equal? Since we have defined a physical theory to be just a (mapping to a) probability distribution, we only need to show that the probability functions given by both theories are the same. We can do that using a *metric* on the space of probability distributions, defined as

$$\mathrm{D}(\mathrm{P},\mathrm{Q}) := \frac{1}{2}\sum_x |\mathrm{P}(x) - \mathrm{Q}(x)| \tag{18}$$

This is a common metric known as the *variational distance*. Using this metric, two distributions P and Q are equal if an only if $\mathrm{D}(\mathrm{P},\mathrm{Q}) = 0$. Another important property of D (and metrics in general) is the *triangle inequality*:

$$\mathrm{D}(\mathrm{P},\mathrm{Q}) \le \mathrm{D}(\mathrm{P},\mathrm{R}) + \mathrm{D}(\mathrm{R},\mathrm{Q}). \tag{19}$$

Both properties follow immediately from definition (18).

Another property that is specific to D is that for two distributions P_1 and P_2 with a joint probability P_{12},

$$\mathrm{D}(\mathrm{P}_1,\mathrm{P}_2) \le \sum_{x\neq y} \mathrm{P}_{12}(x,y). \tag{20}$$

This inequality can be derived with the help of (1):

$$\mathrm{D}(\mathrm{P}_1,\mathrm{P}_2) = \frac{1}{2}\sum_x \left| \sum_y \Big[\mathrm{P}_{12}(x,y) - \mathrm{P}_{12}(y,x) \Big] \right| \le \ldots$$

$$\ldots \le \frac{1}{2}\sum_{x,y} \Big| \mathrm{P}_{12}(x,y) - \mathrm{P}_{12}(y,x) \Big| = \frac{1}{2}\sum_{x\neq y} \Big| \mathrm{P}_{12}(x,y) - \mathrm{P}_{12}(y,x) \Big|$$

Where the terms $\{x = y\}$ drop out of the summation because for those values, $\mathrm{P}_{12}(x,y) - \mathrm{P}_{12}(y,x) = 0$. Finally, for reaching (20), note that

$$\mathrm{P}_{12}(x,y) + \mathrm{P}_{12}(y,x) \ge \max\{\mathrm{P}_{12}(x,y), \mathrm{P}_{12}(y,x)\} \ge \ldots$$

$$\ldots \ge \frac{1}{2}\Big(|\mathrm{P}_{12}(x,y) - \mathrm{P}_{12}(y,x)| + |\mathrm{P}_{12}(y,x) - \mathrm{P}_{12}(x,y)| \Big).$$

This completes the proof.

4.3 Extensibility

Now, it will be shown that about some measurements on some systems, no alternative theory can make better predictions than quantum mechanics. Colbeck and Renner have worked out their argument to be far more general, but we will only reproduce the part that concerns itself with spin measurements on one half of a maximally entangled pair of electrons.

Consider performing a spin measurement on one electron of this pair. Quantum mechanics says that the results of this measurement are distributed uniformly. Define Q to be the uniform distribution on $\{+1, -1\}$, the space of spin measurement results and say that P_1^λ is the distribution of the measurement outcomes given by an alternative theory that is compatible with quantum mechanics as in equation (17). We then have to prove that $\mathrm{Q} = \mathrm{P}_1^\lambda$. This is the case if and only if

$$2\mathrm{D}(\mathrm{P}_1^\lambda, \mathrm{Q}) \stackrel{?}{=} 0; \tag{21}$$

and since for this setup, Q equals the quantum-mechanical distribution P_1^ψ, this is all we have to prove.

Now consider a setup similar to the setup that was used in deriving the chained Bell inequalities in section 3.3: $2N$ measurements, the odd numbered being measurements on the first particle and the even numbered being measurements on the second. Without loss of generality, you can take the first measurement to be the same measurement as the measurement in the previous paragraph. For the other measurements, have the angle between »neighbouring« measurements (that is: measurements with a number 1 apart) be $\pi + \pi/2N$ instead of the $\pi/2N$ that was used earlier. The quantum-mechanical and alternative distributions for these extra measurements will be written as P_n^ψ and P_n^λ.

This setup differs from the original setup of section 3.3 in that compared to the original, this setup has all measurements on the second particle rotated π radians. It leads to violations of the chained Bell inequality that are just as strong, but it does so because of strong correlations instead of strong anticorrelations. It is from these strong correlations that we will show that equation (21) holds.

To make a connection between (21) and the expanded setup, use the fact that Q is the uniform distribution to rewrite

$$2\mathrm{D}(\mathrm{P}_1^\lambda, \mathrm{Q}) = \mathrm{D}(1 - \mathrm{P}_1^\lambda, \mathrm{P}_1^\lambda) \leq \dots \tag{22}$$

Then, you can use the triangle inequality to bring the full array of measurements into the equation:

$$\dots \leq \mathrm{D}(1 - \mathrm{P}_1^\lambda, \mathrm{P}_{(2N)}^\lambda) + \sum_{|a-b|=1} \mathrm{D}(\mathrm{P}_a^\lambda, \mathrm{P}_b^\lambda) \leq \dots \tag{23}$$

where the summation is estimating the distance between P_1^λ and $P_{(2N)}^\lambda$ by summing over the distances between neighbouring measurements.

Equation (20) can be applied to this expression to get rid of the distance function:

$$\ldots \le \sum_x P_{1(2N)}^\lambda(x,x) + \sum_{|a-b|=1} \sum_x P_{ab}^\lambda(-x,x). \tag{24}$$

Note that we were able to sum over $\{x\}$ instead of over $\{x, y \colon x \neq y\}$ because the space of results is just $\{+1, -1\}$. This is responsible for the minus sign inside P_{ab}^λ and the appearance of P_1^λ instead of $(1 - P_1^\lambda)$ in the first term.

Expression (24) looks a bit intimidating, but it can be understood by noting that each of the $2N$ terms is that part of the corresponding correlation function C_{ab} that – in quantum mechanics – nears 0 as $N \to \infty$. The expression as a whole, therefore, contains one half of the terms of the left side of (15). If we had not rotated all measurements on the second particle, we would have ended up with the other terms, the terms that – again, in quantum mechanics – near 1. From this, you can see that we are using the same phenonemon that leads to quantum mechanics violating the chained Bell inequality.

In equation (24), however, we are not using quantum mechanics, but a hypothetical alternative theory. Fortunately, we required that the alternative theory was compatible with quantum mechanics, as expressed in equation (17). This allows us to bring inequality (22–24) into the domain of quantum mechanics:

$$2 \sum_\lambda R(\lambda|\psi) D(P_1^\lambda, Q) \le$$

$$\le \sum_x P_{1(2N)}^\psi(x,x) + \sum_x \sum_{|a-b|=1} P_{1b}^\psi(-x,x) \tag{25}$$

where you should note that we can freely interchange summation signs, because al summations happen to be finite.

Finally, because we know quantum mechanics, we can calculate the value of (25); it is equal to $2N \sin^2 \frac{\pi}{2N}$, which goes to 0 quadratically. And since N could have been chosen arbitrarily high, this gives the result

$$\sum_\lambda R(\lambda|\psi)\, D(P_1^\lambda, Q) = 0. \tag{26}$$

Since neither R nor D can be negative, you can drop the summation to give, for all λ, ψ:

$$R(\lambda|\psi)\, D(P_1^\lambda, Q) = 0. \tag{27}$$

To get from here to (21), choose an arbitrary value for λ. For any value of λ, definition (17) quarantees that there will be at least one compatible quantum state ψ_0 for which $\mathrm{R}(\lambda|\psi_0) > 0$. This means that for $\psi = \psi_0$, you can drop R from the equation, leaving you with the statement that we set out to prove.

4.4 Implications

While the EPR paradox speaks about *elements of reality* or the *simultaneous reality* of complementary observables, the result of Colbeck and Renner is stated in terms of the *extendability of quantum mechanics*. We will now investigate the relations between these two approaches.

As a starting point of our comparison, it is helpful to review the role of the wave function in the discussion so far. Taking the wave function as an object to reason about, the question whether quantum mechanics is complete can be restated as *no quantum state can correspond to multiple ontic states*. This is easy to see by contraposition: if a quantum state can correspond to multiple ontic states, then it can be extended to include those extra states. Conversely, if quantum mechanics can be extended with extra variables to give the ontic state, a quantum state can correspond to multiple ontic states.

In order to prove the incompleteness of quantum mechanics, Einstein, Podolsky and Rosen assumed local realism and a connection between certain measurements and elements of reality. From these assumptions, they established the simultaneous reality of two complementary physical quantities, thus giving two ontic states for one wave function. The Bell inequalities have since shown that the simultaneous reality of complementary physical quantities produces a testable contradiction with quantum mechanics, thereby disqualifying at least one of the assumptions of Einstein, Podolsky and Rosen.

In order to prove the completeness of quantum mechanics, Colbeck and Renner have shown, assuming a form of locality and building on the strong correlations that make quantum mechanics violate the Bell inequalities, that no hidden variables can improve the predictions made about measurement results. Strictly speaking, this does not imply the completeness of quantum mechanics, because it could be the case, for instance, that ontic states are quantum states with an extra parameter that does nothing at all. However, given a set of ontic states for each quantum state, the non-extensibility derived by Colbeck and Renner implies that no information about measurement results will be lost by identifying each set of ontic states with a single state. In this way, it is possible to see quantum mechanics as a complete *model* of reality.

The construction in the previous paragraph of a complete model of reality leaves open the possibility of »merging« two quantum states if the quantum states have overlapping sets of associated ontic states. The question whether that is possible can be formulated as the reverse of the way completeness has been formulated in this section: saying that this »merging« of superfluous quantum states can never occur is equivalent to the statement that *no ontic state can correspond to multiple quantum states*. It is an expression of realism: if an ontic state *can* correspond to multiple quantum states, then the quantum state must contain uncertainty that exists in the mind of the observer.

In an earlier paper,[10] Colbeck and Renner have argued that the non-extensibility of quantum mechanics also implies that no ontic state can have simultaneous existence with two different quantum states. Their argument is rather simple: since by the assumption of non-extensibility, knowledge of the complete, ontic state of a system must lead to the same predictions that could have been made with knowledge of just the wave function of the system, two quantum states that share an ontic state must lead to the same predictions about about the results of *every possible measurement*. In quantum mechanics, this is impossible.

Another perspective on the relation between the quantum-mechanical wave function of a system and its ontic state has been provided by Pusey, Barett and Rudolph.[9] They assume the existence of functions $\mathrm{R}(\lambda|\psi)$ as in (17) and examine whether it is possible for such functions to overlap for different values of ψ. In the terms of the variational distance (18), this can be formulated as

$$\mathrm{D}\Big(\mathrm{R}(\lambda|\psi_0), \mathrm{R}(\lambda|\psi_1)\Big) \stackrel{?}{=} 1 \tag{28}$$

Pusey, Barett and Rudolph show that equation (28) holds if it is possible to prepare independently a large number of copies of ψ_0 and ψ_1, giving a joint state of the form $\psi_{a_1} \otimes \ldots \otimes \psi_{a_n}$ with $a_n \in \{0, 1\}$. This, again, is a non-disturbance condition that follows immediately from locality.

The result of Pusey, Barett and Rudolph is a direct proof of the proposition that no ontic state can correspond to two quantum states. This leaves open the question of quantum mechanics being complete.

5 Conclusions

The questions *»is quantum mechanics a complete theory«* and *»does the quantum-mechanical state express a lack of knowledge of the observer«* can be restated in terms of the relations between the ontic states of reality and the wave functions that are the states of quantum mechanics: if the description of reality given by the quantum state of a system is both complete and a property of the system itself, then quantum-mechanical states and ontic states must be in one-to-one correspondence.

Following the intuïtion of classical mechanics, one might be inclined to think that ontic states must relate to physical quantities. In quantum mechanics this is a problematic idea, because the quantum-mechanical phenomenon of complementary physical quantities seems to contradict their simultaneous reality. This is the background to the EPR paradox: together with the Bell inequalities, the EPR paradox can be seen as a proof that principle of locality is irreconcilable with the idea that a physical quantity corresponds to an element of reality if its value can be predicted with certainty.

Recent publications by Colbeck and Renner[11][10] and Pusey, Barett and Rudolph[9] have shown that some assumptions about locality imply a one-to-one correspondence between ontic and quantum-mechanical states. The argument of Colbeck and Renner depends on the same strong quantum-mechanical correlations that make quantum mechanics violate the Bell inequalities. In both approaches, only models of reality are considered that are compatible with quantum mechanics.

6 References

[1] Einstein, Podolsky and Rosen, *Can quantum-mechanical description of reality really be considered complete?* Physical Review **47/10**, 777–780 (1935)

[2] Bohm, *A suggested interpretation of the quantum theory in terms of »hidden« variables,* Physical Review **85/2**, 166–179 (1952)

[3] Bell, *On the Einstein Podolsky Rosen paradox,* Physics **1/3**, 195–200 (1964)

[4] Clauser, Horne, Shimony and Holt, *Proposed experiment to test local hidden-variable theories*, Physical Review Letters **23/15**, 880–884 (1969)

[5] Fine, *Hidden variables, joint probability and the Bell inequalities,* Physical Review Letters **48/5**, 291–295 (1982)

[6] Garg and Mermin, *Farkas's lemma and the nature of reality: statistical implications of quantum correlations,* Foundations of Physics **14/1**, 1–39 (1984)

[7] Braunstein and Caves, *Wringing out better Bell inequalities,* Annals of Physics **202**, 22–56 (1990)

[8] Hilgevoord e.a, *Grondslagen van de quantummechanica,* in Dutch; lecture notes published by the Institute for History and Foundations of Science, Utrecht University (2004)

[9] Pusey, Barett and Rudolph, *On the reality of the quantum state,* arXiv:1111.3328; Nature Physics **8/6**, 475-478 (2012)

[10] Colbeck and Renner, *Is a system's wave function in one-to-one correspondence with its elements of reality?* arXiv:1111.6597; Physical Review Letters **108/15** 150402 (2012)

[11] Colbeck and Renner, *The completeness of quantum theory for predicting measurement outcomes,* arXiv:1208.4123 (2012)