Universiteit Utrecht

Institute for Theoretical Physics

Master Thesis

The Cellular Automaton and Perspectival Hidden Variable
And how they relate to Bell's Inequalities

Author
P.G.J. Persoon

Supervisors
Prof. dr. G. 't Hooft
Prof. dr. D.G.B.J. Dieks

August 7, 2013

Acknowledgements

My thanks go out to my two supervisors prof.dr.'t Hooft and prof.dr. Dieks, who both took the effort to guide me in this learning process and had to repeatedly break through my stubborn convictions about physics. I am grateful in particular to George Skantzaris and Mathijs Janssen for their help in bringing all this to a good end. Cherique, my parents and my housemates at the Croeselaan deserve a special mentioning for their understanding and support throughout the past year and the time before.

Contents

Chapter 1

Introduction

Quantum Mechanics (QM) has a unique status in terms of acceptance. There is no theory in physics which is as experimentally successful, which is as much debated and about which as many interpretations are around as QM. The fact that problems of interpretation arise can be conceived inherent to the physics enterprise as history teaches, however, these problems generally vanish again at some point. Often this is a matter of changing viewpoints and the introduction of a new ontology, so the questions which were relevant before can be irrelevant after. Yet QM seems to be an exception. The ontology the founding fathers of QM proposed is a collection of statements rather about what we cannot say instead of what we can say about the world [7].

This leads even today, after almost 90 years of QM, to attempts to develop new interpretations and alternatives. We could now explain the situation in three ways: 1) There are nowadays too many physicists and philosophers around who have nothing better to do then discussing old issues. 2) Physics has reached a stage where phenomena are simply too weird to suit any sensible interpretation. 3)QM and its Copenhagen interpretation cannot be the right description of (microscopic) phenomena. Probably most physicists, perhaps for different reasons, would go with the third. In the following work we will be so optimistic to join them in this choice.

Any attempt to developing an alternative to QM has to deal sooner or later with Bell's inequalities. If physical quantities can be thought to be possessed by objects in a local sense and described by a hidden variable, these inequalities place upper bounds to the expectation values of these quantities. QM violates these and seems therefore irreplaceable. We will see in detail for two relatively young attempts, the

Cellular Automaton and the Perspectival Hidden Variable, how the Bell inequalities can or cannot be sidestepped. Roughly said, the Cellular Automaton Model as developed by 't Hooft assumes there to be a reality below the Quantum World which behaves completely deterministically and about which, in principle, the state of affairs can be known in all detail. As the behaviour of this sub-Quantum World is generally quite complicated, QM is used as a convenient means of description. The Perspectival Hidden Variable attempt, which is developed to large extent by this author and about which his supervisors are still doubtful, tries to develop a local description of Quantum Phenomena. This local stochastic hidden variable has the remarkable aspect that it can appear different to observers in different frames of reference.

Our work is structured as follows. In the second chapter it is explained why some consider QM to be strange. By the aid of the hypothetical EPR Bohm experimental set up, we explain how the non-local behaviour arises in the measurements of spin. We see how this behaviour cannot be understood fully using a somewhat naive classical approach to spin. In chapter three, the rigorous proof by Bell is derived which states no local deterministic hidden variable theory could reproduce the same results in this EPR Bohm experiment. It is shown how this also counts when this local hidden variable is stochastic. Then an alternative to QM is discussed in chapter four, which is the Cellular Automaton Model developed by 't Hooft. It is explained how this local deterministic hidden variable theory deals with Bell's Inequalities. Finally in the last chapter the other attempt to an alternative, the Perspectival Hidden Variable, is developed. It is investigated to what extent this model can be made consistent with QM and to what extent it can reproduce the probabilities in the EPR Bohm set up. Then we summarize our results in the chapter with conclusions and finally three appendixes are included which go through some technicalities.

Chapter 2

What is strange about Quantum Mechanics?

In this chapter we will explain the reasons why an alternative to Quantum Mechanics (QM) is desired by some physicists. We will see specifically how the supposed 'strange' non-local effects arise in spin measurements in the EPR Bohm setting. Finally, we will see how a classical approach to spin can ease our minds to some extent with respect to this strangeness, but not fully.

2.1 Quantum Strangeness

It is remarkable to see that on the one hand, physicists can be happy with QM for its great experimental success, whereas on the other hand, so many objections are made to what the theory seems to be telling us about the world. Each formulation (and thereby interpretation) one chooses has a different way in which the supposed weirdness comes out. We will not present a discussion about these different possibilities here (see for example [8]) but will instead highlight the strangeness staying within the Copenhagen Interpretation.

Before the introduction of QM in the 1920s many believed (and many believe still) that events in the world obey deterministic principles. If one would know the laws of physics and the circumstances to infinitely small detail, a prediction can be done giving a 100 percent certainty. In case of wrong or partly wrong prediction, either the wrong laws of physics were used or the circumstances were not known in enough detail. One can choose to ignore information on circumstances and still make predictions, yet those predictions are limited by being statistical. This is, for clarity, what we mean throughout the following with a deterministic perception of the world.

In the Copenhagen Interpretation of QM, this deterministic perception is replaced by an intrinsic probabilistic perception. It is claimed events can generally not be predicted with a 100 percent certainty as a matter of principle. Note that QM by no means claims there is no order in nature, 'stuff still affects stuff' in a predictable way, yet these predictions are formulated in terms of probabilities. The probabilities enter in a way different from the way they enter in the statistical deterministic perception: they are not the result of our ignorance towards available information, but the result of that information being simply non-existent (and therefore not available to us). There is nobody who has that information and nobody who could have that information. One could even go as far as saying the information is non-existent because it is about something non-existent. As this sounds somewhat vague, let us give a clear example: if an observer knows the position of a particle with infinite precision, there is no information at all about the momentum of that particle. One could say it does not even have a certain momentum. That does not mean the particle never has a momentum as a property: once a momentum measurement is done on the particle, it obtains a definite one. Yet the trade off is a total loss of the property position.

Not so much the probabilistic character of QM but this aspect of property loss is what most objections are raised against. Most physicists like the idea that objects in the world have certain properties, whether they are being observed or not. A good theory of physics, they argue, should be able to give some description of these at all times. Yet there is an even more important aspect to these properties which QM totally disrespects. Whether an object has a certain property or not, say in this case it has, it is generally accepted that it can only be affected by something which comes near (so any object at a distance cannot affect it). And if for some reason some action at a distance would affect a property, let it then at least be delayed a time which would be enough for light to cover the distance. (Obeying the principle of Special Relativity that no information can travel faster than the speed of light.)

This idea was formalized by Einstein Podolsky and Rosen (EPR) in a famous paper [2]. It is now considered to be the 'local realist' view. Using as a starting point the requirement that no interaction could act at a speed faster than light and a particular possible experimental set up, they claimed to show QM is a theory which does not give a complete description of reality. According to them, the theory must be supplemented extra variables (hidden variables) which allow 'in principle' to say more about the physical situation. About what these hidden variable should be EPR did not feign any hypotheses, however one can rest assured that they must have done their best finding alternatives.

The EPR paper is often considered the last great offensive and simultaneously (but not at the time) the defeat of the local realist camp. In 1964 J.S. Bell used the same assumptions and a slightly different set up (the EPR Bohm set up, [3]) to derive inequalities any local realist theory should obey [1]. In the next chapter we will see in detail how a derivation goes (as there are many), but let us first fully explain how the EPR set up makes the non local nature of QM manifest. The pedagogical value of this set up to the concept of entanglement is unsurpassed for 78 years now, justifying why we use it as a starting point.

2.2 The EPR Bohm Set Up

In the EPR Bohm set up it is assumed two particles have some interaction for a unspecified time and are then separated. The form and method of this procedure is usually not explained. Often it is imagined to be a spatial separation, which was the case originally for in the EPR paper. What is important, is the assumption that the particles can no longer influence each other in any way. EPR based their derivation largely on the idea that in principle, a measurement 'could be done' on a particle left or right. Bell really required those measurements to be performed on the two particles, each one separately by a different observer. We will call them Alice and Bob. Variations exist as to what is measured, EPR originally used momentum and position. Considering spin instead fairly simplifies calculations because of the finite Hilbert spaces. Both Alice and Bob are assumed to be free in their choice as to along which axis to measure the spin.

Let us consider the specific case of the singlet state, where we have two spin 1/2 particles. As usual, $\vec{S}$ represents the spin operator in the $\vec{n}$ direction as

$$\vec{S} = \frac{\hbar}{2}\vec{n} \cdot \vec{\sigma}, \tag{2.1}$$

where we take $\vec{\sigma}$ to have the Pauli matrices for the corresponding x,y and z components. We consider eigenstates $|s, m\rangle$ of the operators $\vec{S}^2 = (\vec{S}_A + \vec{S}_B)^2$ and $S_z = S_{A,z} + S_{B,z}$, with their respective eigenvalues $\hbar^2 s(s+1)$ and $\hbar m$. The labels A, B signify Alice's and Bobs particle, where $S_{A,z}$ is written for $\sigma_z n_z \otimes I_2$. The singlet

$|0,0\rangle$ is then the eigenstate corresponding to the eigenvalues s and m equal to zero:

$$|0,0\rangle = \frac{1}{\sqrt{2}}\,|z\uparrow\rangle_A \otimes |z\downarrow\rangle_B - \frac{1}{\sqrt{2}}\,|z\downarrow\rangle_A \otimes |z\uparrow\rangle_B\,. \tag{2.2}$$

Here we have written it in the basis of eigenvectors of σ_z. This state is special because of its spherical symmetry: it is similar to the zero eigenvalue eigenstates of S_x and S_y. So any rotation brings it back to itself and the choice for the z-basis is in a sense irrelevant. It is convenient to define the joint, marginal and conditional probability (P^J,P^M,P^C) of a certain (set of) measurements on a state $|\Psi\rangle$ as follows: if $|k_m\rangle_A$ and $|l_n\rangle_B$ are the eigenvectors associated with eigenvalues k_m, l_n,

i) $P^J(k_m, l_n) = |\,\langle k_m \otimes l_n, \Psi\rangle\,|^2$ is the probability Alice measures k_m and Bob l_n,

ii) $P^M(k_m) = \Sigma_n|\,\langle k_m \otimes l_n, \Psi\rangle\,|^2$ is the probability Alice measures k_m regardless of the outcome of Bobs measurement,

iii) $P^C(l_n|k_m) = P_J(k_m, l_n)/P_M(k_m)$ is the probability Bob measures l_n given that Alice measures k_m.

In the EPR set up, anything happening at Alice's side is not supposed to influence Bob's measurement and vice versa. However, this does not mean there cannot be correlations between results of Alice and Bobs. Quite the contrary it is for example in the case of a singlet. If Alice measures spin up along randomly chosen axis z, Bob is sure to measure spin down along the same axis z. Yet the cases: Alice measuring up and Bob down or Alice down and Bob up along the same axes are randomly distributed with a probability of 1/2. So for the EPR singlet set up (see the first appendix), the marginal probability for Alice to measure spin up in the z-direction $P_M(z\uparrow)$ is 1/2, and so is $P_M(z\downarrow)$ and likewise for Bob. The conditional probability $P_C(z_A\uparrow, z_B\uparrow)$ is however $\sin^2[\theta/2]$, where θ is the angle between Alice's axis z_A and Bob's z_B. The moment Alice does a spin measurement, the probabilities on Bobs side are promptly affected depending on the way she chose her axis. Because both Alice and Bob are free to choose their axes until the very last moment, it is as if there is an immediate two-way responsiveness between the particles.

The two-way responsiveness is in terms of probabilities. For a good part, we can reason away the strangeness comparing the situation to a classical deterministic story. Say there is a vase with a black and a with ball in it. Alice and Bob can each draw one ball blindly and at random from the vase, put it a suitcase and travel to different ends of the world. Each have a probability 1/2 of having the black ball in their suitcase. Yet if Alice opens her suitcase and observes her ball is the white one, Bob's probability for measuring black has 'suddenly jumped' from 1/2 to 1. If instead Bob had opened his suitcase it would have been Alice's probability which changed. So in a sense, one can expect such a immediate two way responsiveness classically. Yet we can never fully reason away the strangeness. In QM, it turns out this responsiveness is far stronger than one would expect classically. To make this more quantitative let us compare the case of the singlet to a possible classical description of the same situation.

2.3 Classical Spin Model

We largely base the following discussion of on the approach presented in [4]. Classically, the spin of an object is pictured as a rotation around an axis of symmetry. It is one of the two forms of angular momentum, the other being a orbital angular momentum, which is a rotation around any other axis. The spin of a particle is then represented by a vector $\vec{L} = I\vec{\omega}$, where I is the moment of inertia and $\vec{\omega}$ the angular velocity. It is clear I can have any value, and if one would like to measure a component of L along a certain axis z (directed along the unit vector $\vec{z}$), this value will be reduced by a factor $\vec{\omega}\cdot\vec{z}$, creating a continuous spectrum of $\vec{L}$ along different z axes.

This vector picture of spin breaks down more or less in quantum mechanical spin situations (where we start considering the spin of very small objects). For example from the Stern-Gerlach experiment, we know that the spin measurements of little dipoles by no means presents a continuous spectrum of values along different z axes. Along each axis, only two possible values are distinguished. In the Stern Gerlach experiment these values are distinguishable by an upward or downward deflection of a beam of magnetic dipole atoms going through an inhomogeneous magnetic field, so we can call them 'up' and 'down'. In a quantitative sense, they can be represented by $+1$ and -1. If these values should be reproduced as components of a classical spin vector L along a vector $\vec{z}$, it is clear the values must be normalized in some sense. So for the spin component Alice measures along her axis of choice z_A:

$$s_A = [1, -1] = \frac{\vec{L}\cdot\vec{z_A}}{|\vec{L}\cdot\vec{z_A}|}, \tag{2.3}$$

so the sign of the spin flips if the axis $\vec{z_A}$ makes an angle greater than $\pi/2$ with the spin vector $\vec{L}$. We see that the choice for $\vec{z_A}$ defines a 'northern hemisphere' on the sphere defined by all other choices for $\vec{z_A}$. The value for s_A is $+1$ if the tip of the vector $\vec{L}$ is somewhere in this northern hemisphere and -1 otherwise (See figure).

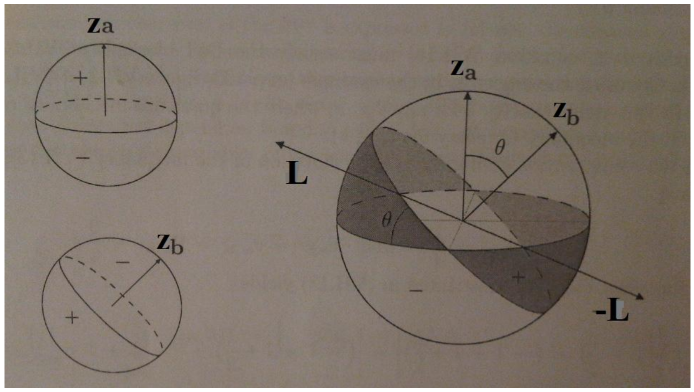

Figure 2.1: Classical Spin Vectors (adjusted from [4]).

Now we will try to adopt this model as an alternative to the EPR Bohm set up with a singlet state. We then have two particles where the values [+1,−1] are perfectly anti-correlated if from both particles the spin measurement is along the same axis. So then we can deduce for Bob we get a spin component along his axis of choice z_B:

$$s_B = [1, -1] = -\frac{\vec{L} \cdot \vec{z_B}}{|\vec{L} \cdot \vec{z_B}|}. \tag{2.4}$$

For Bob's particle too we can define a northern and southern hemisphere defined by his choice for $\vec{z_B}$. Placing the two spheres over one another defines four areas on the sphere as illustrated in the included figure. We are interested in the area of the two shaded parts of the surface together and of the remaining parts taken together. θ_{z_A,z_B} is the angle between $\vec{z_A}$ and $\vec{z_B}$. The area of the shaded surface is then $4\theta_{z_A,z_B}$, of the remaining part it is $4(\pi - \theta_{z_A,z_B})$. Let us now assume the distribution of different $\vec{L}$ is isotropic, which says the amount of particle pairs with the tip of $\vec{L}$ in a certain area of the sphere is for a large number of particle pairs the same as the amount in any other certain area of the sphere, where the size of the two areas are arbitrary but equal. So the number of particle pairs having the tip of $\vec{L}$ in a certain area of the sphere is proportional to size of the area. The size of the shaded area is the proportional to the fraction of cases where the value of the product of s_A and s_B is positive, for the remaining area this quantity is negative. For a given choice of z_A, z_B we can then calculate the average of the product of s_A and s_B by

$$\langle s_A s_B \rangle = \frac{1}{4\pi}\big(\theta_{z_A,z_B} - 4(\pi - \theta_{z_A,z_B})\big) = -1 + \frac{2}{\pi}\theta_{z_A,z_B}, \tag{2.5}$$

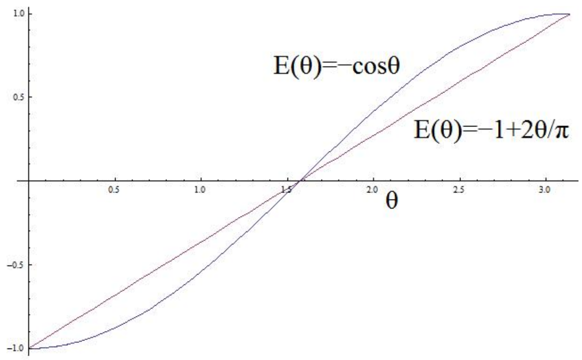

Figure 2.2: A plot of the expectation values.

which is plotted in the following figure together with a $-\cos\theta_{z_A,z_B}$. The cosine is what QM predicts to be the $\langle s_A s_B\rangle$ of the singlet and it is clear its absolute value is convincingly greater than the absolute value derived in the classical model for all values of $\theta_{z_A z_B}$, except for the -1,0 and $+1$ expectation values (for $\theta_{z_A z_B} = 0, \pi/2, \pi$). It turns out this was not just a bad choice for a classical model similar to the EPR singlet situation (actually it is quite a decent approximation as we will see in the next chapter). It will follow that whatever classical local model one chooses, the correlations are always weaker than QM would predict. Bell delivered the proof for this statement in 1964 [1], we will discuss it in the next chapter.

Chapter 3

Bell Inequalities

We will now see in detail how Bell first derived certain quantitative statements which any local realist theory should satisfy and which Quantum Mechanics violates. We will largely follow a pedagogical derivation based on [4]. In the first section we will consider local deterministic hidden variables and in the second local stochastic hidden variables. As was mentioned in the previous chapter Bell uses the same set up as in the EPR paper but then slightly adjusted to the case of two spin 1/2 particles as Bohm proposed. So it amounts to doing experiments on two different particles, separated far enough, so no interaction between the two can influence measurement outcomes. The experimenters, Alice and Bob, can choose their axes z_A and z_B along which they measure the spin value of their particle freely. We should be precise what we mean with free choice, this discussion we reserve for the third section of this chapter. Throughout this chapter it is supposed we are dealing with the singlet state again, so a perfect (anti)-correlation for the spins can be identified for any direction if Alice chooses the same axis as Bob.

3.1 Local deterministic hidden variables

The perfect (anti)-correlation of the singlet state allows one to predict the spin value of Bob's particle with a 100 percent certainty if the spin value of Alice's particle is known. Any local realist hidden variable model should then take the spin value to be a property which the particle 'possesses' in a local sense. This is in accordance with the definition EPR gave for an 'element of reality' [2]. One could see the spin value as a physical aspect of the particle which can only be influenced by something which comes near. A local realist theory of physics should be ambitious enough to be able to predict this spin value for a given pair of particles using a hidden variable λ. Let us denote the value of the spin if it is up by $+1$ and for down -1. Most generally speaking, these spin values of Alice's particle s_A and Bob's s_B would then depend on λ,z_A and z_B. Yet because the spin property of Alice's particle cannot be affected non-locally, it should not matter if Bob changes his mind and instead chooses an axis z'_B. Therefore $s_A = s_A(\lambda, z_A)$ and $s_B = s_B(\lambda, z_B)$. This is where one could object, saying that if the whole experimental set up obeys a causal (deterministic) structure, it is determined which axes will be chosen. So Alice and Bob cannot change their minds, or if they would, the spin values would be a function of their actual choice. We will discuss this at the end of this chapter.

It is then clear that for the singlet, we have for any choice of axis z_A

$$s_A(z_A, \lambda) = -s_B(z_A, \lambda). \tag{3.1}$$

Contrary to QM, where a state $|\Psi\rangle$ is taken to be identical every trial of the experiment, a local realist model distinguishes different possibilities $|\Psi(\lambda_1)\rangle$ and $|\Psi(\lambda_2)\rangle$ (and more) each giving a specific measurement outcome. In this derivation, it is supposed not much is known about the apparatus producing the different $|\Psi(\lambda)\rangle$s. Instead some probability density $\rho(\lambda)$ is assumed which does not depend on z_A and z_B. This makes sense, as after producing a certain $|\Psi(\lambda)\rangle$, Alice and Bob could still change their minds about which axis to measure against. (Again this is where objection may be raised.) Assuming a continuous λ over a domain Λ, we require a normalization

$$\int_\Lambda \rho(\lambda) d\lambda = 1. \tag{3.2}$$

So we can deduce that the expectation value $E(z_A, z_B)$ of the product of s_A and s_B is then (we no longer use the $\langle\rangle$ notation),

$$E(z_A, z_B) = \int_\Lambda s_A(\lambda, z_A) s_B(\lambda, z_B) \rho(\lambda) d\lambda, \tag{3.3}$$

and therefore, using (3.1),

$$E(z_A, z_B) = -\int_\Lambda s_A(\lambda, z_A) s_A(\lambda, z_B) \rho(\lambda) d\lambda. \tag{3.4}$$

Obviously, $s_A(z_A, \lambda)^2 = 1$. So for another choice of Bob for z'_B

$$\begin{aligned} E(z_A, z_B) - E(z_A, z'_B) &= -\int_\Lambda (s_A(\lambda, z_A) s_A(\lambda, z_B) - s_A(\lambda, z_A) s_A(\lambda, z'_B)) \rho(\lambda) d\lambda, \\ &= \int_\Lambda s_A(\lambda, z_A) s_A(\lambda, z_B)(s_A(\lambda, z_B) s_A(\lambda, z'_B) - 1) \rho(\lambda) d\lambda. \end{aligned}$$

Taking the absolute values left and right, using that $|s_A(\lambda, z_A)s_A(\lambda, z_B)| = 1$, it follows,

$$\begin{aligned} |E(z_A, z_B) - E(z_A, z'_B)| &\leq \int_\Lambda (s_A(\lambda, z_B)s_A(\lambda, z'_B) - 1)\rho(\lambda)d\lambda, \\ |E(z_A, z_B) - E(z_A, z'_B)| &\leq E_{AB}(z_B, z'_B) + 1. \end{aligned} \tag{3.5}$$

This inequality is the original Bell inequality, however there are many more variations one can think of. Starting for example from

$$E(z'_A, z_B) + E(z'_A, z'_B) = \int_\Lambda (s_A(\lambda, z'_A)s_A(\lambda, z_B) - s_A(\lambda, z'_A)s_A(\lambda, z'_B))\rho(\lambda)d\lambda \tag{3.6}$$

We obtain going through the same steps that

$$|E(z'_A, z_B) + E(z'_A, z'_B)| \leq 1 - E_{AB}(z_B, z'_B). \tag{3.7}$$

This can then be added to (3.5) to obtain:

$$|E(z_A, z_B) - E(z_A, z'_B)| + |E(z'_A, z_B) + E(z'_A, z'_B)| \leq 2. \tag{3.8}$$

This inequality is known as the Clauser, Horne, Shimony and Holt (CHSH) inequality, derived using slightly different requirements in [5]. Whereas any local realist hidden variable theory respecting the assumptions made should obey these inequalities, we will see now that QM violates them for some choices for the axes. The expectation value $E(z_A, z_B)$ according to QM is given by,

$$E_{QM}(z_A, z_B) = -\cos\theta, \tag{3.9}$$

where θ is the angle between z_A and z_B. Choosing all z axes to be in the same plane, $z'_A = z_B$ with which z_A and z'_B make angles θ and $-\theta$ respectively (see figure for this configuration), we obtain for the expression on the left hand side of the CHSH inequality:

$$|-\cos\theta + \cos 2\theta| + |-\cos\theta + 1|, \tag{3.10}$$

which is greater than 2 for all $\theta < \pi/2$. We can now use this result to 'test' the classical spin theory which was discussed at the end of chapter 1. It was based on the idea that the spin is a vector like quantity. For the EPR Bohm singlet set up, it was imagined there were many arbitrary orientated particle pairs with a perfect anti correlation left and right. If we similarly to the case considered previously choose all

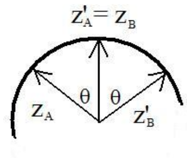

Figure 3.1: A particular configuration of axes

z axes to be in the same plane, $z'_A = z_B$ with which z_A and z'_B make angles θ and $-\theta$ respectively, we find for the expression on the left hand side of the CHSH inequality:

$$|(-1+\frac{2}{\pi}\theta)-(-1+\frac{4}{\pi}\theta)|+|-1+(-1+\frac{2}{\pi}\theta)| \tag{3.11}$$

which equals 2 (for all $\theta \leq \pi$), so it perfectly obeys the inequality for all θ. As a matter of fact, we see that the classical spin model produces the maximum attainable value, hence our suggestion this model is a fairly good approximation. We saw that QM on the other hand violates the inequality for all θ.

There is no ambiguity about how to interpret the inequality, nor about how QM violates it. It is up to nature as the absolute authority to decide what relation the expectation values obey. As of today experiments showed a confirmation of QM [6]. The conclusion must be that if there is a hidden variable theory which could reproduce the same results as QM, it would have to be a non-local theory.

3.2 Local stochastic hidden variables

As QM and rivalling hidden variable theories often deal with processes on the smallest workable space and time scales any precise observation of 'what is going on' is problematic if not impossible. Only after many repetitions of an experiment, where the circumstances are assumed to be identical every time, statements about probabilities can be derived. QM in its usual Copenhagen interpretation can only predict these probabilities and is not secretive about it. Perhaps hidden variable theories should not be too ambitious either and be content with a stochastic hidden variable. Bell showed that his theorem also applies to these cases. We take our approach again largely from [4] where we stay within the EPR Bohm set up. There are some demands which we want our stochastic hidden variable theory to meet.

(i) Parameter independence: The conditional probability for Alice to measure s_A

given a certain value λ is independent of Bob's choice for z_B,

$$P^C_{z_A,z_B}(s_A|\lambda) = P^C_{z_A}(s_A|\lambda). \tag{3.12}$$

(ii) Outcome independence: The conditional probability for Alice to measure a s_A only depends on the choices for the axes z_A, z_B and the stochastic hidden variable λ. So to be explicit, it does not depend on s_B,

$$P^C_{z_A,z_B}(s_A|s_B,\lambda) = P^C_{z_A,z_B}(s_A|\lambda). \tag{3.13}$$

(iii) Source independence: The statistical distribution $\rho_{z_A z_B}(\lambda)$ of λ as produced by the source does not depend on Alice and Bob's choice for z_A and z_B,

$$\rho_{z_A z_B}(\lambda) = \rho(\lambda). \tag{3.14}$$

These requirements of course count for Bob as well. We will now show how the CHSH inequality can be derived from these assumptions. First we write down in a very general sense the joint probability $P^J_{z_A,z_B}(\lambda, s_A, s_B)$ for Alice and Bob to measure s_A and s_B,

$$P^J_{z_A,z_B}(\lambda, s_A, s_B) = P^C_{z_A,z_B}(s_A|s_B,\lambda)P^C_{z_A,z_B}(s_B|\lambda)\rho_{z_A z_B}(\lambda). \tag{3.15}$$

Here $P^C_{z_A,z_B}(s_A|s_B,\lambda)$ is the conditional probability for Alice to measure up if it is given Bob measures s_B and for a given value λ) and $P^C_{z_A,z_B}(s_B|\lambda)$ is the conditional probability for Bob to measure s_B for a given value of λ. ($\rho_{z_A z_B}(\lambda)$ is the probability to have a value λ in this situation. Using the demands, we see this greatly simplifies to

$$P^J_{z_A,z_B}(\lambda, s_A, s_B) = P^C_{z_A}(s_A|\lambda)P^C_{z_B}(s_B|\lambda)\rho(\lambda). \tag{3.16}$$

Now we are ready to calculate the expectation value $E(z_A, z_B)$ of s_A times s_B. We simply multiply the joint probability for a specific outcome with the specific values (s_A and s_B are $[+1,-1]$), obtaining

$$E(z_A, z_B) = \int_\Lambda \left(P^J_{z_A,z_B}(\lambda,1,1) - P^J_{z_A,z_B}(\lambda,1,-1) - P^J_{z_A,z_B}(\lambda,-1,1) + P^J_{z_A,z_B}(\lambda,-1,-1)\right)d\lambda. \tag{3.17}$$

Using previous factorization and sorting terms, we obtain

$$E(z_A, z_B) = \int_\Lambda \Big(P^C_{z_A}(1|\lambda) - P^C_{z_A}(-1|\lambda)\Big)\Big(P^C_{z_B}(1|\lambda) - P^C_{z_B}(-1|\lambda)\Big)\rho(\lambda)\lambda. \tag{3.18}$$

We note that the terms

$$|P^C_{z_A}(1|\lambda) - P^C_{z_A}(-1|\lambda)| = |f_{z_A}(\lambda)| \leq 1, \tag{3.19}$$
$$|P^C_{z_B}(1|\lambda) - P^C_{z_B}(-1|\lambda)| = |g_{z_B}(\lambda)| \leq 1. \tag{3.20}$$

So now a difference or sum between two expectation values can be written as

$$E(z_A, z_B) - E(z_A, z'_B) = \int_\Lambda f_{z_A}(\lambda)\big(g_{z_B}(\lambda) - g_{z'_B}(\lambda)\big)\rho(\lambda)d\lambda, \tag{3.21}$$
$$|E(z_A, z_B) - E(z_A, z'_B)| \leq \int_\Lambda |f_{z_A}(\lambda)||g_{z_B}(\lambda) - g_{z'_B}(\lambda)|\rho(\lambda)d\lambda. \tag{3.22}$$

Using (3.19),

$$|E(z_A, z_B) - E(z_A, z'_B)| \leq \int_\Lambda |g_{z_B}(\lambda) - g_{z'_B}(\lambda)|\rho(\lambda)d\lambda. \tag{3.23}$$

Likewise one can derive

$$|E(z'_A, z_B) + E(z'_A, z'_B)| \leq \int_\Lambda |g_{z_B}(\lambda) + g_{z'_B}(\lambda)|\rho(\lambda)d\lambda. \tag{3.24}$$

Adding these, using that for any z, $|x| \leq 1$ and $|y| \leq 1$,

$$\int_\Lambda |g_z(\lambda)|\rho(\lambda)d\lambda \leq 1 \qquad |x+y| + |x-y| \leq 2 \tag{3.25}$$

So with the intermediate conclusion,

$$\int_\Lambda (|g_{z_B}(\lambda) + g_{z'_B}(\lambda)| + |g_{z_B}(\lambda) - g_{z'_B}(\lambda)|)\rho(\lambda)d\lambda \leq 2, \tag{3.26}$$

we arrive at the CHSH inequality for a stochastic hidden variable:

$$|E(z'_A, z_B) + E(z'_A, z'_B)| + |E(z_A, z_B) - E(z_A, z'_B)| \leq 2. \tag{3.27}$$

3.3 Free Will

In the derivations an assumption of free will was used. It means that we allow both Alice and Bob to 'change their mind' as to along which axis to measure the spin against at the very last moment. This is essential for the assumption of source independence. Somebody holding on to a deterministic (classical) view could object.

In a world which is deterministic, even the actions of the experimenters are predetermined and so, Alice and Bob cannot change their mind at the last moment, and if they do, this was predetermined as well. So in principle, the information about which axes are chosen is available and stored in some way in the system, so it could be the particles are affected by it. This is of course a perfectly valid argument, and we will see how it enables one to justify alternatives to QM. Yet the argument may not be satisfactory to everybody. The complexity of how Alice's choice of axes is affected by what she had for breakfast two weeks before the experiment is a problem far beyond the ambition of physics, yet a deterministic world view would claim it relevant for the prediction of the outcome. In that sense the problem is dismissed to be one of physics. Also, a new problem arises, because one would have to explain why, if it is averaged over many trials which were considered different, such a precise prediction as that of QM is obeyed. Was not the charm of the enterprise as well to be able to give a fairly good description of reality having as little information available? How much empirical success should an ugly ontology offer in order to replace one that is beautiful? We will continue this discussion at the end of next chapter.

Chapter 4

The Cellular Automaton Model

It may be strange to start considering local realist alternatives to Quantum Mechanics (QM) only after we derived that they can never reproduce the same predictions as QM. Yet it is interesting to see how theories manage or fail to produce that extra piece of prediction. In this chapter we will discuss an approach proposed by 't Hooft where microscopic processes are pictured as deterministic interactions between very small but simple cells. For a short history of this approach see [11].

Young as the following approach is, clearly the interpretation of QM is not a closed case at the time of this writing. In [9] 't Hooft relies on a classical world view from which the quantum world emerges in a new way. QM, it is claimed, is only a powerful tool to describe processes which are fundamentally classical. In a more picturesque sense, 'below' the world of quantum phenomena (more or less around the 10^{-18} meter scale), there is a world (imagined to be somewhere around the Planck scale 10^{-35} meter) which is again suitable for a classical description. This means that classical principles (or as 't Hooft prefers, 'logic') apply to it in the same way as it applies to macroscopic phenomena. We will refer to it here as the sub-quantum world.

With the principles of a classical world it is meant mainly the rule that processes evolve deterministically. For every given initial state of affairs, the next state can be predicted unambiguously (and therefore also the one after that). Note however, this does not mean that the reverse counts as well. In the models proposed by 't Hooft, it does not necessarily count the past can be unambiguously derived from the present. Room is left open for so called 'information loss' or 'memory loss'. The description of the sub-quantum world is in terms of discrete time and space. This allows for the simple interpretation of every point in space 'processing a bit of information' at

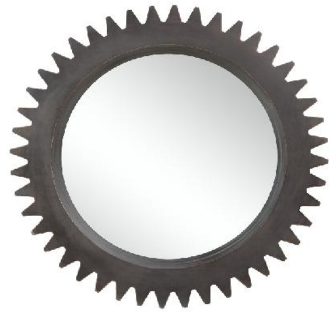

Figure 4.1: A typical Cogwheel.

every point in time. To make these ideas somewhat more tangible let us discuss a baby example: the Cogwheel Model.

4.1 The Cogwheel Model

Imagine a cogwheel with N teeth rotating over an angle $2\pi/N$ for every fundamental time step Δt (which, logically, we choose to be the value of our time unit). Obviously this would be a model for a cycle repeating itself infinitely.

We have the states $|0\rangle$,$|1\rangle$, ..., $|N-1\rangle$ each corresponding to a certain orientation of the wheel, with $|0\rangle$ corresponding to the situation at $t = 0$. The evolution law is not hard to define, for $0 \leq n \leq N-2$:

$$t \to t+1 \qquad |n\rangle \to |n+1\rangle \, , \tag{4.1}$$

and for $n = N-1$

$$|n\rangle \to |0\rangle \, . \tag{4.2}$$

One can barely think of a simpler model but it very well serves the purpose to explain how we can now introduce QM as a tool to do statistics. The states $|n\rangle$ can be interpreted as an orthonormal basis of a finite dimensional Hilbert space. A general quantum state $|\Psi(t)\rangle$ can then be written

$$|\Psi(t)\rangle = \sum_{n=0}^{N-1} \alpha_n(t) \, |n\rangle \, , \tag{4.3}$$

where the coefficients $\alpha_n(t)$ can be interpreted as amplitudes, leading to the probability $|\alpha_n(t)|^2$ to be in state $|n\rangle$ at time t. Normally we define the evolution of such a state by a unitary operator U,

$$|\Psi(t+1)\rangle = U\,|\Psi(t)\rangle \tag{4.4}$$

Where U is given

$$U = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 1 & \ddots & & 0 \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{pmatrix}$$

So that for the first one time step

$$\begin{pmatrix} \alpha_0(0) \\ \alpha_1(0 \\ \vdots \\ \alpha_{N-1}(0) \end{pmatrix} \rightarrow \begin{pmatrix} \alpha_{N-1}(0) \\ \alpha_0(0) \\ \vdots \\ \alpha_{N-2}(0) \end{pmatrix}$$

So

$$\begin{aligned} \alpha_0(0) &= \alpha_{N-1}(1) \\ \alpha_1(0) &= \alpha_0(1) \\ \alpha_2(0) &= \alpha_1(1) \\ &\dots \\ \alpha_{N-1}(0) &= \alpha_{N-2}(1) \end{aligned}$$

The coefficients of $|\Psi(t)\rangle$ for other t follow likewise. This illustrates how the probabilities to detect a certain state hop over time to other states. So we have done nothing more than describing the time evolution in a quantum mechanical way. Of course nothing stops us from going to a different basis as is usual in QM. If we choose it such that U is diagonal,

$$U = \begin{pmatrix} 1 & 0 & \dots & \\ 0 & e^{-2\pi i/N} & & \\ \vdots & & \ddots & \\ & & & e^{-2\pi i(N-1)/N} \end{pmatrix}$$

We can write $U = e^{-iH}$, so we identify an Hamiltonian

$$H = \frac{2\pi}{N} \begin{pmatrix} 0 & 0 & \dots & \\ 0 & 1 & & \\ \vdots & & \ddots & \\ & & & N-1 \end{pmatrix}.$$

We see that we obtain eigenvalues which resemble the first couple of natural numbers. Note the similarity, up to a constant, with the energy levels of the harmonic oscillator in QM. The Cogwheel Model can also be described in terms of a continuous variable t. The wheel is then supposed to have an infinite number of teeth. Again there is an interesting analogy identifiable between this model and the harmonic oscillator in QM. This already hints there are systems in QM which could perfectly well be described by a deterministic systems (but for which the QM description is just very convenient).

4.2 Beables, Changeables and Superimposables

't Hooft stresses that an important distinction needs to be made between the operators which are diagonal in the original (ontological) basis and those which are not. The first type he calls Beables, corresponding to eigenstates which represent the actual state of affairs in the world. All Beables for a given time t commute. The second type of operators are either Changeables or Superimposables. Changeables, amongst other for example U, turn one ontological state into another. The third type, Superimposable, mixes multiple ontological states into a superposition. (Not to be confused with what is usually called a mixed state in QM). Superimposables are most generic operators. 't Hooft considers Changeables and Superimposables not to have an ontological interpretation as the Beables have.

In QM it happens though, that Beables sometimes evolve into Changeables. This is for example the case for a particle with a certain spin which starts rotating in a magnetic field. In the quantum mechanical description of this process, if the spin is measured along a z axis, the operator σ_z can be taken to be a Beable. Yet when a magnetic field is applied along the y axis, the operator σ_z can be interpreted to rotate to become the operator σ_x. Yet this operator σ_x was taken to be a Changeable when viewed from the (eigen) basis of σ_z. 't Hooft argues that a theory which allows for a state which has an ontological interpretation to evolve into a state which does not have one, cannot be a deterministic theory. When a deterministic interpretation is desired, this could be considered the deeper reason why QM is neither fundamental nor complete. An underlying theory should be able to explain why Beables, Changeables and Superimposables get so thoroughly mixed up in our quantum description of reality. 't Hooft suspects this is the result of a renormalization group transformation, but we will not go into that. The observables in this underlying theory, describing processes at the Planckian level, should all be Beables and hence they will only evolve into Beables. As soon as larger scales are considered, somehow the structure of the quantum theory should appear with its Changeables and Superimposables. The following model may just have that property.

4.3 The Cellular Automaton

The Cogwheel model was just a taste of the structure which 't Hooft proposes as a alternative to QM. At the heart of his interpretation is a more advanced (yet simple) model: the Cellular Automaton. It attempts to be an alternative to a Quantum Field Theory. It has a flavour of the simulations often run on computers (to simulate a biological process for example) where a bunch of cells with permanent location evolve through time on the basis of a set of simple rules. These rules are often such that only the neighbouring cells are relevant to the development of a single cell. This gives the dynamics of the system a very local character. Although these rules are often simple, the evolution and dynamics of the system as a whole may become very complicated.

In the physical description of the Cellular Automaton (CA) as 't Hooft proposes, both time and space are discrete. To not lose any generality a D-dimensional space lattice is introduced, so any position can be indicated by a vector $\vec{x} = (x_1, x_{2,...,x_D})$ where all x_n are integers. The evolution of the whole system could drawn with full precision in a comic book, a window for every step in time $\Delta t = 1$. The physical variable $F(\vec{x}, t)$ could represent anything, we take it to be a certain integer modulo some natural number N. It is convenient for the formulation of the evolution law to attach these physical degrees of freedom to even sites in spacetime, so for

$$\sum_{i=1}^{D} x_i + t = even. \tag{4.5}$$

See the next figure to envisage such a system for $D = 3$.

The recursive relation which describes the dynamics is the equivalent of the rules in computer simulations and can be given in a general form, for $\sum_i x_i + t = odd$ by

$$F(\vec{x}, t+1) = F(\vec{x}, t-1) + \\ Q\big(F(x_1 \pm 1, x_2, ..., x_D, t), F(x_1, x_2 \pm 1, ..., x_D, t), ..., F(x_1, x_2, ..., x_D \pm 1, t)\big) \text{ Mod N}, \tag{4.6}$$

So this is the physical value, which is an integer, of cell c at $\vec{x}$ at time $t+1$ (future). Q is some function of the indicated integers, which are the physical values of the

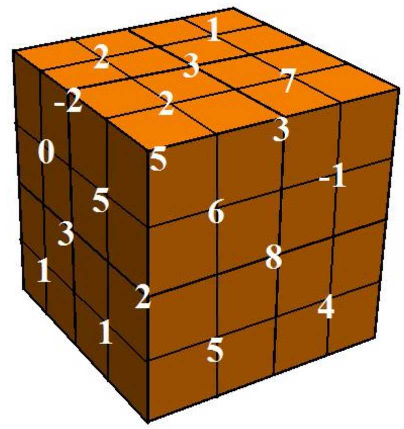

Figure 4.2: A cellular Automaton for $D = 3$.

neighbouring cells of c at t (present) and the cell c at $t - 1$ (past). Note it is not a function of the physical value of the cell c at t as we only defined physical values at even spacetime lattice sites and this is not the case for c at t. Note also that the time evolution of this system is completely time reversible: from the set of data for two subsequent steps in time one can work both forward and backward in time. The specific values of the integers at every point in space basically consist of one state of the ontological basis. Together with a evolution law, this would be enough to describe the universe.

So far for the ontological states. We will now see how this model would be described quantum mechanically. It is first switched to the Schrödinger picture, where the states instead of the operators are time dependent. So $F(\vec{x}, t) \to F$ and we split it up in a part $X(\vec{x})$ which acts on the even sites and a part $Y(\vec{x})$ which acts on the uneven sites. Of course $X(\vec{x})$ and $Y(\vec{x})$ need to be 'updated' alternately every t. That is why we choose the evolution operator $U(t, t-2)$ to be defined over two time steps, where it is defined for even t that

$$U(t, t-2) = AB, \tag{4.7}$$

where A updates $X(\vec{x})$ and B updates $Y(\vec{x})$. The operators A and B can defined further as

$$A = \prod_{\vec{x}=even} A(\vec{x}) \qquad B = \prod_{\vec{x}=odd} B(\vec{x}). \tag{4.8}$$

They obey the commutation relations

$$[A(\vec{x}), A(\vec{x'})] = 0 \qquad [B(\vec{x}), B(\vec{x'})] = 0. \tag{4.9}$$

Yet if $|\vec{x} - \vec{x'}| = 1$, (so if $\vec{x}$ and $\vec{x'}$ are neighbours)

$$[A(\vec{x}), B(\vec{x'})] \neq 0. \tag{4.10}$$

All other A's and B's do commute. The operators act in finite dimensional subspaces of the total Hilbert space, so we can write $A(\vec{x}) = e^{-ia(\vec{x})}$ and $B(\vec{x}) = e^{-ib(\vec{x})}$, where

$$a(\vec{x}) = -P_x(\vec{x})Q(\{Y\}) \qquad b(\vec{x}) = -P_y(\vec{x})Q(\{X\}). \tag{4.11}$$

The $P_x(\vec{x})$ is the generator of a one step displacement of $X(\vec{x})$ modulo N

$$e^{iP_x(\vec{x})}\left|X(\vec{x})\right\rangle = \left|X(\vec{x}) + 1 \quad \text{Mod N}\right\rangle, \tag{4.12}$$

and similar for the generator P_y on the values Y. It must be that for all $\vec{x}$ and $\vec{x'}$:

$$[a(\vec{x}), a(\vec{x'})] = 0 \qquad [b(\vec{x}), b(\vec{x'})] = 0, \tag{4.13}$$

and for $|\vec{x} - \vec{x}| > 1$,

$$[a(\vec{x}), b(\vec{x'})] = 0. \tag{4.14}$$

As these operators commute we can rewrite

$$A = exp[-i \sum_{\vec{x}=even} a(\vec{x})] \qquad B = exp[-i \sum_{\vec{x}=odd} b(\vec{x})]. \tag{4.15}$$

We are now ready to calculate the evolution operator U for two subsequent time steps

$$U = \exp[-i \sum_{\vec{x}=even} a(\vec{x})] \exp[-i \sum_{\vec{x}=odd} b(\vec{x})] = e^{-2iH}. \tag{4.16}$$

This can only be calculated using the Campbell-Baker-Hausdorff formula:

$$\begin{aligned} &e^S e^T = e^R \\ &R = S + T + \frac{1}{2}[S,T] + \frac{1}{12}[S,[S,T]] + \frac{1}{12}[[S,T],T]] + \frac{1}{24}[[S,[S,T]],T] + ..., \end{aligned} \tag{4.17}$$

of which we assume it converges. It follows the Hamiltonian can be written as a sum of Hamiltonian densities

$$H = \sum_{\vec{x}} \mathcal{H}(\vec{x}), \tag{4.18}$$

$$\mathcal{H}(\vec{x}) = \frac{a(\vec{x})}{2} + \frac{b(\vec{x})}{2} + \mathcal{H}_2(\vec{x}) + \mathcal{H}_3(\vec{x}) + ... \tag{4.19}$$

Where

$$\mathcal{H}_2(\vec{x}) = \frac{-i}{4} \sum_{\vec{y} \text{ neighbours of } \vec{x}} [a(\vec{x}), b(\vec{y})], \tag{4.20}$$

$$\mathcal{H}_3(\vec{x}) = \frac{-1}{24} \sum_{\vec{y}_1, \vec{y}_2 \text{ neighbours of } \vec{x}} [a(\vec{x}) - b(\vec{x}), [a(\vec{y}_1), b(\vec{y}_2)]], \tag{4.21}$$

The terms become quickly very complicated. Next to nearest neighbours interaction terms enter for higher order terms. Yet we can draw the conclusion that for $|\vec{x}-\vec{x}'| \gg 1$,

$$[\mathcal{H}(\vec{x}), \mathcal{H}(\vec{x}')] = 0. \tag{4.22}$$

At any finite order of the series the Hamilton density is $\mathcal{H}$ is a finite matrix. Its lowest eigenvalue can be set equal to Planck's constant h. Therefore, for any large but finite volume V, the total Hamiltonian H also has a lowest eigenvalue $E_0 > hV$. The eigenstate corresponding to this eigenvalue is then usually called the vacuum $|0\rangle$. So note that although this state is stationary in a QM language, it does not mean there is nothing happening to the automata. One can then go the next eigenvalue, of which the eigenstate would describe a single particle state, etc. See how a Quantum Field Theory is given shape here. The states arising this way should be subject to any orthodox or Copenhagen interpretation of QM.

Of course this result was rather sketchy and we sort of stepped over three major differences with any conventional Quantum Field Theory. First of all, neither Galilean nor Lorentz invariance nor rotational invariance is incorporated in the theory. There are extensions of the theory which manage to incorporate invariance, see [12]. The evolution law, keeps open the possibility that the right neighbouring cell has a different influence on a cell c than a left neighbour would. In the next section this is further investigated. It should be investigated if the particles which arise obey any dispersion relations which in some way resemble those which we know. It is not hard to see that we will never be in trouble going faster than a certain speed (of light): as the 'real' interactions between the cells are only local, a piece of information maximally takes one step in space per time unit. This brings us to the second difference, which is the notion that space and time would be fundamentally discrete. Although maybe hard to believe, it is not deemed impossible such very small quantities exist. Third, as was already mentioned, it was assumed that the Campbell-Baker-Hausdorff formula converges. This is a delicate issue and requires more investigation. In [10] a conjecture is presented that it will converge if it is sandwiched between to energy eigenstates $|E_1\rangle$, $|E_2\rangle$ with corresponding eigenvalues E_1, E_2 such that $2|E_2 - E_1| < 2\pi\hbar/\Delta t$, where Δt is one step in time, which we take to be the Planck time. This is a large restriction, but it may turn out that this energy appearing in the exponent of (4.16) is really al lot more than the energy typically considered in QM. Yet this is a conjecture so again, caution is required.

4.4 Evolution law considerations

Without any further specification of the evolution law, the conclusions we can draw are limited. One could try to work towards a Quantum Field Theory as we know it, building in principles such as Lorentz invariance and only allowing degrees of freedom as we know them from the Standard Model. The precise form of the evolution law can then be constructed to meet the imposed requirements. 't Hooft has made several steps in this direction [10]. In the following we take a different approach. Instead we impose some a priori 'reasonable' characteristics on the cellular automata and see how this further restricts the evolution law and possibly the QM description of the automata. This may seem a rather ambitious approach, as assuming determinism in the first place was a dared enough starting point. Yet any new theory should have a more or less independent ontology from which the physical principles follow, so let us try and see what results can be obtained.

What would be a reasonable characteristic for a cellular automaton? A cell should be an object as simple as possible, having a physical value F with N degrees of freedom. The only complicated aspect to a cell is that it feels in some way the physical values of its neighbours and is as a function of these affected two steps in time later. As we require the cells to be as simple as possible, it seems acceptable they would not know the difference between left and right (up and down, etc). The influence of the left neighbouring cell should be similar to that of the right neighbouring cell, in the sense that if the physical values of the left and right neighbours are interchanged, the effect on the cell should be similar. This should then be imposed for any direction $x, y, ..etc$, as no direction should be special in this sense.

If the physical values of left and right can be interchanged at will their contribution to the Q term in the evolution law consists on two levels: 1) An independent contribution of left F_L and right F_R, which should be similar: $f(F)$. 2) A contribution $g(f(F_L), f(F_R))$ of these together, in which $f(F_L)$ and $f(F_R)$ can be interchanged. An elementary way of satisfying this requirement is to assume g is simply a function of the sum $f(F_L) + f(F_R)$. One could argue to take a product instead, but then if they are written in an exponential form the sum is simply retrieved (at the cost of taking $f'(F) = Log[f(F)]$ instead). So we obtain for $\sum_i x_i + t = odd$:

$$\begin{gathered} F(\vec{x}, t+1) = F(\vec{x}, t-1) + \\ Q\Big(g_1\big(f_1(F(x_1+1, x_2, ..., x_D, t) + f_1(F(x_1-1, x_2, ..., x_D, t)\big), \\ g_2\big(f_2(F(x_1, x_2+1, ..., x_D, t)) + f_2(F(x_1, x_2-1, ..., x_D, t)\big), ... \\ g_D\big(f_D(F(x_1, x_2, ..., x_D+1, t)) + f_D(F(x_1, x_2, ..., x_D-1, t))\big)\Big) \text{ Mod N.} \end{gathered} \tag{4.23}$$

In a way this is an assumption of parity for an individual cell. This does not mean that there is also parity in the QM description of this system (the parity we are familiar with), as this applies to particles/constituents/objects on a different scale, which are represented by a whole collection of cellular automata. A collection of these automata not necessarily obey a parity symmetry.

We could go further by imposing this symmetry and also assume the cell cannot distinguish the different directions. So if left&right would be interchanged by up&down, we require the evolution law to have the same effect. We can then rewrite Q as a function of the sum of gs, which may be a transformed version of the previous gs, for odd $\sum_i x_i + t$

$$\begin{gathered} F(\vec{x}, t+1) = F(\vec{x}, t-1) + \\ Q\Big(g_1\big(f_1(F(x_1+1, x_2, ..., x_D, t)) + f_1(F(x_1-1, x_2, ..., x_D, t))\big) + \\ g_2\big(f_2(F(x_1, x_2+1, ..., x_D, t)) + f_2(F(x_1, x_2-1, ..., x_D, t)\big) + ... \\ g_D\big(f_D(F(x_1, x_2, ..., x_D+1, t)) + f_D(F(x_1, x_2, ..., x_D-1, t)\big)\Big) \text{ Mod N} \end{gathered} \tag{4.24}$$

Note that it should count that

$$\begin{gathered} g_n\big(f_n(F(x_n+1, x_m)) + f_n(F(x_n-1, x_m)\big) + g_m\big(f_m(F(x_n, x_m+1)) + f_m(F(x_n, x_m-1))\big) = \\ g_n\big(f_n(F(x_n, x_m+1)) + f_n(F(x_n, x_m-1))\big) + g_m\big(f_m(F(x_n+1, x_m)) + f_m(F(x_n-1, x_m))\big) \end{gathered} \tag{4.25}$$

for all n, m and x_n and x_m (we did not write the dependence on other x_k). We can therefore conclude that all g and f are similar for all n, giving

$$\begin{aligned} F(\vec{x}, t+1) = F(\vec{x}, t-1) + \\ Q\Big(g\big(f(F(x_1+1, x_2, ..., x_D, t)) + f(F(x_1-1, x_2, ..., x_D, t))\big) + \\ g\big(f(F(x_1, x_2+1, ..., x_D, t)) + f(F(x_1, x_2-1, ..., x_D, t))\big) + ... \\ g\big(f(F(x_1, x_2, ..., x_D+1, t)) + f(F(x_1, x_2, ..., x_D-1, t))\big)\Big) \text{ Mod N} \end{aligned} \tag{4.26}$$

In a way we could again interpret this assumption as some discrete rotational invariance (for 90 degrees angles) for an individual cell. Likewise to the case of parity, this does not mean that in the QM description there is also the familiar rotational invariance.

4.5 Relation Cellular Automata and Bell inequalities

4.5.1 Superdeterminism versus independence

Any hidden variable theory is at some point confronted with the Bell inequalities. The hidden variable of the Cellular Automaton is what was defined as the ontological basis. If one knows the 'real state of affairs' for only two instances of time, on can 'play forward' the film at will to see how history develops. For systems without any memory loss, one can also rewind the film at will. The course of history can in principle be developed (or retrieved for no memory loss) with exact precision, until one ends up at some beginning, supposedly the Big Bang. If a measurable quantity A is said to be correlated with B in some experiment, then in every repetition of the experiment, some specific values of A imply to some extent some specific values of B (or vice versa). Often some causal relation can explain this responsiveness, but this does not necessarily mean A causes B: it could be both A and B have a certain common cause C in the past.

It may be clear that for the Cellular Automata model without memory loss, although the interactions are local, the behaviour of any two cells a and b is correlated, as can be seen if 'the film' is only rewound far enough, there is always a common cause

C at some point in the past. It would therefore a priori be a mistake to suppose independence in the statistical behaviour of two distant events. The idea is then that of course b can be correlated with a in a way such that statistical expectations are exceeded, as the statistical independence is only an illusion. In the world we perceive, everything is correlated to everything to a dazzling precision. This point of view is often expressed as Super Determinism. As 't Hooft would say: "If you believe in determinism, you have to believe in it all the way." An aspect which is often overlooked in these discussions is that the initial state of an experiment as EPR Bohm can be far from trivial. It is often assumed the three different macroscopically separated parties, surrounded by a vacuum, act independently, because no significant effect is expected from a vacuum. Yet if one had to define an initial state for an EPR Bohm experiment in a Cellular Automata model, where even the vacuum evolves non trivially, the structure of the correlation can be very well encoded in this initial state and it should hence be no surprise these correlation are demonstrated when the experiment is performed.

In a theoretical sense, the Bell inequalities do not apply to Cellular Automata for (1) the assumption of source independence and (2) the assumption that a spin value on Bob's side does not depend on Alice's choice of axes. Breaking (1) is enough to sidestep the inequalities. If $\rho_{z_B,z_A}(\lambda)$ would be the probability for the source to give a particle pair of which the spins are described by the hidden variable λ, remember that it was assumed to be independent of z_A and z_B. In a Super Deterministic approach, one would not make such an assumption. It follows that, in an attempt to derive a Bell inequality by constructing a sum of expectation values $E(z_A, z_B), E(z_A, z_B)...$, rho_{z_A,z_B} no longer factors out and the derivation is stuck.

The Bell inequalities would not be so popular if this would be the end of the story. Determinism is an intuitive principle, yet it is also intuitive idea that some systems are (to some extent) causally independent. There is some scientific value in saying all systems in our universe are in some way (causally) connected, yet the approach of seeing systems as separate entities proves to be successful too. The idea that all hydrogen atoms work more or less the same is quite a dared and powerful statement. As soon as we separate a system from its environment, a simplification is implied, which tells us the ultimate result can never be 'completely correct'. However, it seems inevitable if a scientist wants to construct a theory from some hypothetical objects, these represent something of a more or less physically independent nature 'out there'.

Next to being independent, it is generally desired that these objects have a limited number of degrees of freedom. In other words, that a rather simple description would be sufficient to describe all possible behaviour of the system.

Combining these two points, an isolated system (so under no influence of an environment) is supposed to go through a development in time which is specific to that system. If the situation would be repeatedly investigated on similar systems, a similar development is expected. Say this could be backed up by experiments. It would then be a small step to conclude that, after these isolated systems went through their characteristic developments, they have 'forgotten' all about their previous connections with environments. It is then sort of assumed that if systems possess different memories, they would also show different behaviour. It is therefore tempting to conclude if they show similar behaviour, they either have similar memories or no memory at all.

To say this somewhat more clearly: it is hard to believe that two given water molecules A and B each from a different sea of the world, showing similar behaviour in almost all experiments we know, still have in some way encoded the enormous story of their lifetime (and possibly more than that). It could even be argued the right scientific approach should be the opposite: to try and reduce nature as far as possible to these independent yet similar objects and give them as little properties as possible.

4.5.2 Memory Capacity

For the theory of Cellular Automata too it would be reasonable to ask how far back a certain cause can be retrieved for it to still make a meaningful difference at present. In a sense we would ask how good some constituents of the system function as memory keeping devices: a memory capacity. For the case where the evolution of the Automata is 100 percent deterministic, so that the integer value of a given cell in the past can be retrieved with no uncertainty, the memory is of the system is perfect: there is no boundary on how far back history can be retrieved. Let us suppose there would be some creature which can operate in a world outside the cells and can, for some time $t = 0$, adjust the integer value of some cell instantly. Then a time $t = n$ later, all cells a number of steps n further can feel the difference, no matter how large n.

It seems therefore that the Bell inequalities can be rather easily circumvented. The assumption of Alice and Bob to have a free will is not valid, as the outcome of a spin measurement at Bob's side is intricately correlated with Alice's choice of axes. The assumption that these are statistically independent is therefore unjustified. In a world with Cellular Automata, one can therefore expect very non-trivial correlations to occur when some classical independence would be expected. If the Cellular Automata Model should be an alternative to QM, the EPR Bohm experimental set up would be a situation where this can be directly observed. The Cellular Automata Model would then automatically suffer from the same problem QM deals with: how is it then possible 99 percent of the world we perceive behaves according to some classical rules in which independence seems manifest?

For Cellular Automata with only very little memory loss, the situation becomes completely different. Memory loss comes about when the future follows unambiguously from the past but the past does not follow unambiguously from the future. 't Hooft keeps the possibility open that the evolution laws for some Cellular Automata models show indeed some memory loss, although its role is not exactly clear yet. If memory loss should apply to the universe, it can even be seen as an elegant feature: possibly a great number of initial states led to the fact that some order was created (planets, human beings, etc).

There is multiple ways to talk about memory loss and to investigate its effects. Often it is explained as a tendency of a system in which there is variety to evolve into a system with less variety. In terms of Cellular Automata, it enforces homogeneity of the physical values of multiple cells. Yet we would like to focus on the effect of memory loss for an individual cell and less on the effect on a large scale. Here we propose it to mean the following: for a cell c with memory loss together with the set of its neighbouring cells and their physical values at some time t, it cannot be concluded with complete certainty what the physical value of c was at $t-1$. Say we have a system described as the above, where each cell has N degrees of freedom and a complicated evolution law. The only difference is that we now impose the weakest memory loss imaginable. Say c has $2D$ neighbours. Then only when we have configuration M, which says that at time t, c has specific value m_0, and the neighbouring cells $x_1, x_2, x_3...$ have specific values $m_1, m_2, m_3...$ we say it cannot be decided whether c had a value m_a or m_b at $t-1$. It could be we attribute a probability p and $1-p$ to these possibilities. As often when nothing indicates that the one case is more likely than the other, the principle of indifference suggests to take both to be 1/2. Yet if the memory loss should be kept to a minimum, one of those probabilities should become very small.

We cannot determine the value of the cell c at $t-1$ given the configuration M as a matter of principle, not the cell c, nor any observer knowing the evolution law can derive it with absolute certainty, hence the name memory loss. The form we introduced above is really the weakest form of memory loss (with the meaning we give it here) imaginable for the Cellular Automata Model, as is clear from the many ways in which it can be extended: we could introduce more configurations M_1, M_2... each having their own ambiguity. We could introduce weaker configurations which are sufficient for an ambiguity: only a part of the neighbouring cells would then need to have certain values. We could extend the size of an individual ambiguity, so then one would consider a number > 2 of possibilities when a specific configuration M occurs. Finally, as was already remarked, the probabilities of these possibilities could be increased to make a certain ambiguity more likely. (So for k possibilities for a given ambiguity, $1/k$ for each of the possibilities would be the case of maximum memory loss.)

Can anything quantitative be said about independence of distant physical objects given this little memory loss? For this purpose we should define independence more precisely. First of all, dependence (or 'correlation') occurs when two physical values m_A and m_B of two distant cells A and B have a common cause in the past. Say the number of cells between them is $2n+1$, where n to be even for simplicity, then the common cause is at least a space time distance of $n+1$ away. If the cells would know about their unique individual history, say characterized by some sequence m_1, m_2, m_3..., they could with certainty detect an adjustment if instead of m_3 they suddenly take the value m_3'.

For the model without memory loss, we saw that the arbitrary adjustment of one cell by a creature would, for big enough n, affect all possible cells at some point in time. This is no different for a system with memory loss (as the future follows unambiguously from the past), however, this does not mean that from a given certain physical value of a distant cell, it can be said with certainty some adjustments were made in the past. If in a specific chain of interaction some ambiguous values occur, different possible histories are identifiable. Some of these histories may contain an adjustment of the creature, some do not. Of course, even in a model without memory loss, a distant cell some steps n further could only feel the adjustment some time n later. So we see that for any cell c, we can define a specific distance s outside which a possible adjustment is detectable with a probability smaller than $1/2$, a minimal time $t > s$ later. Said differently, if there is an adjustment outside s, it is more likely to go unnoticed than to be detected.

We are now ready to define independence: Two cells a and b are said to be independent if the number of cells n between them is greater than two times the characteristic distance s. In that way, a common cause in the past a time $t > s$ ago could be adjusted, and the likelihood for a and b to detect it would be smaller than 1/2. It does not make sense to speak of a common cause in the past if it could have been just as well another cause, hence a and b are called independent.

If we could find out a typical value for the distance s it would be a rough measure for the memory capacity of the system. In order to avoid the Bell inequalities, a Cellular Automata model should enable the memory capacity to be of macroscopic order. Only then it makes sense to view Alice choice of axes and the spin of Bobs particle not to be independent.

However, it seems without any knowledge of the evolution law, nothing can be concluded about the size of this distance s. It could even be the evolution law 'conspires' to avoid ambiguities. In that case our analysis would end here, so let us assume this is not the case. We would like to approximate the evolution law to be more quantitative about the memory capacity. Of course it would be too crude of an assumption that for a given time t, a given cell c is equally likely to have physical value m as any other value m'. Yet any improvement on this approximation would ask for more information on the physics (the evolution law), which we do not have for now. For a system where the cells have N degrees of freedom, the probability to have a value m_0 would then be $1/N$. The probability to have configuration M, so for all its neighbouring cells to have those specific values $m_1, m_2, m_3...$ would be $1/N^{2d}$. (Note that if rotational invariance would be imposed, this probability would fairly increase.) So a rough estimate for c and its neighbours to have a configuration M would be $1/N^{2d+1}$. Say given the configuration M is present, there is still a probability to have an ambiguity of p. So then we obtain the probability to have an ambiguity of p/N^{2d+1}

At every time t, c and its neighbours can basically be ambiguous or not. The probability for not being ambiguous for a number of n time steps is $(1 - p/N^{2d+1})^n$. This probability is smaller than 1/2 for

$$(1 - \frac{p}{N^{2d+1}}) \leq 2^{-\frac{1}{n}}, \tag{4.27}$$

$$n \leq \frac{\log[\frac{1}{2}]}{\log[1 - \frac{p}{N^{2d+1}}]}. \tag{4.28}$$

Let us now suppose an adjustment a time t ago cannot be excluded when a single ambiguity arises. The time steps of the Cellular Automata model are generally very small, as they should appear infinitesimal from a macroscopic point of view. According to 't Hooft, they must be of the Planck scale, which is of the order 10^{-44} second. So say we wish to investigate what kind of Cellular Automaton with a least imaginable memory loss could satisfy the above relation for macroscopic times, so $n = 10^{44}$, then

$$e^{\frac{\log[1/2]}{10^{44}}} \leq 1 - \frac{p}{N^{2d+1}}, \tag{4.29}$$

$$\frac{p}{N^{2d+1}} \leq \frac{\log[2]}{10^{44}}. \tag{4.30}$$

Let us consider the possibility to have an ambiguity given the configuration M to be quite unlikely, so p is very small $\approx 10^{-k}$, for k some natural number > 1. Then still we would require $N \approx 10^{43-2d-k}$ to have cells which have an average unambiguous memory for about a second. More precisely, we should say the probability for them to have an unambiguous memory after one second is on average less than $1/2$. It may be clear that we need N to be enormous in order to stay above this probability. It could be that the number of the degrees of freedom are actually continuous, but then it would be unfair to still induce an ambiguity which only appears at very special configurations. The probability for having this special configuration is then practically zero and it is really dealt with a system with no memory loss. It could also be the number of dimensions are very large, say 35. This has as a disadvantage that every cell has 70 nearest neighbours, making the evolution law depend on 70 variables. It may be very hard to design an interaction which is more or less simple but yet depending on 70 variables.

Let us go back to the EPR Bohm set up. Alice and Bob are typically a macroscopic distance apart. In a rough approximation, we have seen that the Cellular Automata model with only a little memory loss would need to have a very strange form if Alice's choice and the spin of Bob's particle should still be dependent. It would be necessary for them to be dependent in order to avoid the Bell inequalities.

Maybe we went over some assumptions somewhat too quick in the previous approximation. First, it was assumed a single ambiguity in a the chain of interactions over time would be enough for an adjustment to become undetectable. This assumption merely simplifies the approach, as if for example two ambiguities would be necessary, one has to sum over all possible ways in which a single ambiguity can occur in a given chain of interactions. This result in a probability

$$n(p_a/N^{2d+1})(1-p_a/N^{2d+1})^{n-1}+(1-p_a/N^{2d+1})^n \approx (np_a/N^{2d+1}+1)(1-p_a/N^{2d+1})^n \tag{4.31}$$

Although the factor n in front may look impressive it is complicated with something which is almost its multiplicative inverse (up to a couple of order ten). The resulting pre-factor would still no comparison to the n in the exponent.

Second, maybe the adjustment of just one cell by the creature is too small of an adjustment: to make a real difference the creature should change a couple of cells next to one another. This would definitely complicate the approximation but we shall not investigate it here, as it requires far more assumptions on the evolution law. As a response we could say this is more or less cheating. The content of a common cause for A and B can be 'summarized' in the value of the single cell to which both A and B have an equal, least spacetime distance. The value summarizes the history where the past 'light cones' of A and B overlap. All other cells are in the causal past of either A or B (or of neither) and have nothing to do with possible dependence between A and B.

Third, maybe the assumption all physical values the cell can take on are equally probable is unjustified. It could be there is a range of physical values which cells only rarely take on. If the vacuum around us should be represented by cells, this seems probable at first (not much is going on). Yet QM tells us that on quantum scales, the vacuum is a big chaos of spontaneous particle creation and annihilation. Second thoughts therefore seem to indicate cells do take on the values of their full spectrum occasionally. These remarks however are merely speculative.

In a very general sense, the problem with the Cellular Automata model, if it should be a serious alternative to QM, is that this ontological basis is unknown to us and probably very complicated. We have no idea how the evolution law for the cells is specifically defined. For a model without memory loss it seems obvious the Bell inequalities could be violated, yet it is not clear by how much and in what way. For the case of some memory loss, the Cellular Automaton seems to be more attackable in terms of independence. Yet these are very rough approximations still.

Chapter 5

Perspectival Hidden Variable

The Bell inequalities can be taken as a guide towards designing a hidden variable theory. To be a serious alternative to QM, a theory had better violate the inequalities or at least avoid them in a certain way. The following approach may provide an unexpected way to sidestep Bell's inequalities. To a local realist, any weird local hidden variable is better than no local hidden variable, so why should we not introduce a hidden variable of which the value is observer dependent. To illustrate what we mean with an observer dependence, we shortly discuss how it arises in Einstein's Theory of Special Relativity (SR). Next we will show how the Bell inequality is sidestepped and finally we will see how such a hidden variable could determine probabilities.

5.1 Analogy between PHV and SR

In SR an observer dependence is introduced. This means observers in different reference frames can disagree on the value of certain physical quantities, namely, time and space intervals. More precise: if we would have two observers Alice and Bob each in possession of an identical clock, then a difference in velocity $\vec{v}$ between them results in the fact that Alice sees Bob's clock ticking slow (compared to her own clock) by a factor

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{5.1}$$

where $\vec{c}$ is the velocity of light. Because of the symmetry of this situation, of course Bob claims Alice is the one moving, so her clock is ticking slow by the factor γ. To our knowledge there is no experiment which could prove one of them right. Of course, if

Alice and Bob knew their SR, they would understand there is really no disagreement if both statements are taken in a different frame. (Needless to say, if both make effort to go to the same frame, all 'disagreements' vanish). However, that does not mean the clocks only seem to tick slower (in the way, for example a stick in bucket of water seems to be kinked at the water surface). Time dilation is a physical effect. If Alice would accelerate her clock to the frame of Bob, leave it there for a while, and then decelerate it back to her frame of reference, it will indeed show the clock has aged less than another clock of Alice which had been left untouched during the process. So in a sense, there is a meaningful way for Alice and Bob to decide which one of them clocks ticks slower: from the common knowledge who of them previously accelerated in the past to establish the difference. Yet this is only an agreement, based on history.

A Perspectival Hidden Variable (PHV) is assumed to be observer dependent. We also allow ourselves the freedom to make it a stochastic variable (it can always be investigated later if a deterministic version is possible too). We will work out the PHV hypothesis for the EPR Bohm set up. If the value of the PHV depends on the frame of the observer then what would characterize such a frame in the EPR Bohm set up? It would be good if unlike the variable, which is hidden, the frame can be considered as something tangible (a feature of our classical world). In the following we shall try to make the idea of a reference frame precise.

As the PHV alternative needs to be compared to the QM approach, it is convenient to define a reference frame using the terminology of QM. We again consider a situation where two observers, let us dub them again Alice and Bob, do different measurements on a general pure quantum state. What we mean here by doing measurements on quantum states effectively amounts to doing a certain experiment many times and thereby obtaining a probability. QM assumes the quantum state is then identical each time the experiment takes place, whereas a PHV approach would not. For now we will consider situations described by states which are bipartite, meaning, its Hilbert space $\mathscr{H}$ can be factorized in two smaller ones $\mathscr{H}_1 \otimes \mathscr{H}_2$ such that the states in both Hilbert spaces can be operated upon separately (using, for example $\vec{S} \otimes I$ or $I \otimes \vec{S}$) with operators corresponding to observables. Any measurement on a bipartite pure state defines two frames of reference, one corresponding to a measurement of an observable of an operator acting in $\mathscr{H}_1$, the other in $\mathscr{H}_2$.

The 'settings' of the these measurements can be chosen independently. For the EPR Bohm case with two spin 1/2 particles for example, Alice can choose to measure the spin along a different axis then Bob. A reference frame can therefore be characterised by the particular choice of setting of the measurement device. (Whoever from whatever angle reads the result from the device we take to be irrelevant, we take this to be an interaction which can be perfectly described by classical physics). It is therefore fair to say that if two observers choose the same settings for their measurement apparatus, they are in the same reference frame. Of course, there is no special, absolute '0 degree' orientation. Yet in the same way Alice claims to be at rest in her frame and no dilation or contraction factor apply to objects in her frame in SR, she will claim her orientation is the '0 degree' one and the PVH has its 'genuine value'. Bob claims the same in a different frame. How would this enable us to sidestep Bell's inequalities?

5.2 Sidestepping Bell's Inequality

In the derivation of the CHSH inequality for a stochastic variable (see (3.27)), it was assumed Alice and Bob conceived of the same value for λ. In the PHV approach, Alice supposes that Bob measures of a different value λ' than she does, depending on which frame he is in. In order to still do some physics, let us suppose there is some relation between the two

$$\lambda' = \Lambda(\lambda, \vec{z_A}, \vec{z_B}), \tag{5.2}$$

$$\lambda = \Lambda^{-1}(\lambda', \vec{z_A}, \vec{z_B}), \tag{5.3}$$

for some function or operator Λ and where $\vec{z_A}$ and $\vec{z_B}$ are the directions of the axes z_A and z_B along which Alice and Bob choose to measure the spin. So we suppose that every time Alice would measure λ, she knows Bob will measure a λ'. One of the requirements for a local stochastic hidden variable theory should be parameter independence (see chapter 2). For the PHV case, it is seems the requirement is only satisfied partly. Alice claims the conditional probability $P^C_{z_A}(s_A|\lambda)$ for her to measure to measure a certain value s_A is independent of Bob's choice for z_B, so that is good. On the other hand, to her, Bob's conditional probability to measure a value s_B is $P^{C,A}_{z_B}(s_B|\lambda)$, which is calculated on the basis of the transformed values $\lambda = \Lambda^{-1}(\lambda', \vec{z_A}, \vec{z_B})$. As the transformation of the λs depends on the axes, the probability of Bob as Alice perceives it does depend on her choice for z_A, so it would be fair to write $P^{C,A}_{z_B}(s_B|\lambda, z_A)$. Bob sees this all different. Bob would claim his conditional probability $P^C_{z_B}(s_B|\lambda)$ is unaffected by Alice choice, instead it is her probability $P^{C,B}_{z_A}(s_A|\lambda_1, z_B)$ which depends on his choice of axis. We learn from SR not to ask the question who is right in these matters. Both observe the right physics in their own frame.

However, we would like to argue the parameter independence requirement is not just satisfied partly. The requirement of parameter independence normally blocks two observers from exchanging signals instantaneously, as an 'actio in distans' (as argued in citation lecture note foundation of QM). With an ordinary hidden variable λ_0 in a theory not satisfying parameter independence, such a thing would be possible. Bob could choose a direction z_B determining a certain probability for Alice to measure s_A. Alice could instantaneously, knowing the λ_0 states emitted by the source (which are a bunch each time), determine the relative frequency and thereby the probability for s_A. Message sent.

A PHV does not allow such a structure, because the probability for Alice to measure s_A as perceived by Bob is generally not the same as the value she measures in her frame (we will see how this works in the next section). To make the argument clear intuitively, let us use the analogy with SR again, (which we take to be a parameter independent theory). Say Alice sees Bob's clock tick slow by a factor $\gamma(\vec{v})$ because he is moving with a velocity $\vec{v}$. Then she boosts herself along the same direction to a certain velocity $\vec{w}$ and sees that after that, Bob's clock instead ticks slow by a factor $\gamma(\vec{v} \oplus \vec{w})$ (where $\oplus$ represents relativistic addition of parallel velocities). There is no experiment Alice can do in her new frame which will have a different result in her old frame, so to her, she is still at rest. Now, of course, the physics of Bob's clock were not affected at all as a result of Alice changing frames (actio in distans). It is only the physics of Bob's clock as perceived by Alice which changed. Likewise, say that Alice would have a λ meter in her frame, then if she would change frame, she assumes her λ meter still works fine in the new frame. She will take the λ_a before she changed frame as 'genuine' a value as the value λ_A after she changed frame. Say Bob measures a λ_b in his frame. Before she changed frame, Alice would say Bob really measures a $\Lambda^{-1}(\lambda_b) = \lambda_a$. After she changed frame this changes. On the basis of their new orientations (no matter how they ended up that way), the transformation Λ becomes a different transformation Λ'. So Alice would claim this time Bob really measures $\Lambda'^{-1}(\lambda_b) = \lambda_A$. Alice's perception of λ in Bob's frame from before she changed frame differs from the value after she changed. Yet Bob, throughout the whole time, just perceives the same λ_b on his meter. Ultimately, it would be a matter of definition whether or not parameter independence is satisfied or not. Let us at least be clear that actio in distans is prohibited.

To the derivation of the Bell inequality (or actually, the CHSH inequality), previous considerations are crucial. Following the steps explained in chapter 2 section

2, the expression for the expectation value, as calculated from Alice frame, becomes

$$E(z_A, z_B) = \int_\Lambda \left(P^C_{z_A}(1|\lambda) - P^C_{z_A}(-1|\lambda)\right)\left(P^{C,A}_{z_B}(1|\lambda, z_A) - P^{C,A}_{z_B}(-1|\lambda, z_A)\right)\rho(\lambda)d\lambda. \quad (5.4)$$

Again defining

$$|P^C_{z_A}(1|\lambda) - P^C_{z_A}(-1|\lambda)| = |f_{z_A}(\lambda)| \leq 1, \quad (5.5)$$

$$|P^{C,A}_{z_B}(1|\lambda, z_A) - P^{C,A}_{z_B}(-1|\lambda, z_A)| = |g_{z_B}(\lambda, z_A)| \leq 1, \quad (5.6)$$

we see that it can generally NOT be concluded

$$\int_\Lambda (|g_{z_B}(\lambda, z'_A) + g_{z'_B}(\lambda, z'_A)| + |g_{z_B}(\lambda, z_A) - g_{z'_B}(\lambda, z_A)|)\rho(\lambda)d\lambda \leq 2. \quad (5.7)$$

As $|x + y| + |x' - y'|$ is not necessarily smaller than 2 for all $|x|, |x'|, |y|, |y'| \leq 1$.

For now we cannot say how λ' depends on λ and the axes, so we cannot say more about this expectation value as calculated by Alice. However, the calculation done from Bob's frame had better agree with that result, because ultimately, an experiment will provide a unique expectation value averaging over the PHV. (Alice and Bob would not want to argue over the validity of a list of spin $\uparrow$ and $\downarrow$.) If Bob would predict a different expectation value this would go against the equivalence of their reference frames.

5.3 How λ determines probabilities

In experiments done so far the predictions of QM have been confirmed with an astonishing precision. It does not matter whether it is Alice or Bob who processes a list of data produced by an experiment: both will determine the same marginal and conditional probabilities, both for Alice's and Bob's measurements. So far, it is not clear how in the PHV approach, Alice can cook up a different value for the probability for Bob to measure s_B than he actually measures in his frame if she would use the same data. How can this be explained? This is where the relevance of λ comes in. QM, in its usual interpretation, assumes the state $|\Psi\rangle$ to be identical every repetition of the experiment. The PHV approach does not assume such a thing: it divides the data in subsets for different λ's: λ_1,λ_2,λ_3,... Let us now suppose our variable λ is not so hidden any more, meaning that both Alice and Bob can have a certain λ meter identifying two possible outcomes λ_1 and λ_2. We imagine it to be some measurable aspect of the particle (or something coming along with the particle) at both sides.

5.3.1 Similar frames

It is important to emphasize that we assume here that if Alice and Bob would be in the same frame, they would measure the same value of λ with a probability of 1. This is required for it to be a hidden variable, because if it would only be specific to the individual particle, it would be no different from just another observable of the particle commuting with its spin. EPR would attribute an element of reality to λ according to their definition in (citation). We demand λ to encode some information about both the particle left and right. It could be that the λ measured at Alice's side is different from the λ measured at Bob's side when they are in the same frame, but then there should at least be a relation which gives the value at Bob's side with certainty given Alice's value. So the assumption we made really says that this relation is a trivial equality, to keep it simple. Therefore, if Alice and Bob would be in the same frame, choosing the direction of their axes similar, a possible table with outcomes of EPR Bohm experiment, for some general bipartite state, could be

trial	Alice measures along z_A		Bob measures along $z_B = z_A$	
	spin value	λ_A	spin value	λ_B
1	$\downarrow$	λ_1	$\uparrow$	λ_1
2	$\uparrow$	λ_2	$\uparrow$	λ_2
3	$\uparrow$	λ_1	$\downarrow$	λ_1
4	$\downarrow$	λ_1	$\uparrow$	λ_1
5	$\uparrow$	λ_2	$\uparrow$	λ_2
etc				

Ultimately, the fact whether Alice measures a λ_1 or λ_2 should help her determine the probability to measure spin up or down, as any good stochastic hidden variable should do. This is implemented by introducing the (conditional) probability $P^C_{z_A}(\uparrow|\lambda_1)$ for her to measure $\uparrow$ for a given λ_1. To calculate it, Alice simply needs to separate in her λ_A data the λ_1 and λ_2 cases. Likewise Bob determines a $P^C_{z_B}(\uparrow|\lambda_1)$ on the basis of his subset for λ_1, which is similar to Alice's in this case. Let us say the size of Alice's subset defined by λ_1 is $N_{A,1}$ and so $N_{A,2} = N - N_{A,1}$. Likewise the size of the Bob's subset defined by λ_1 is $N_{B,1}$. So for Alice and Bob in the same frame we have $N_{A,1} = N_{B,1}$ and $N - N_{A,2} = N - N_{B,2}$. The number of $\uparrow$'s in the subset defined by λ_1 we call $n_{\uparrow,A,1}$. Alice determines (with Bob's agreement, as he sees the same values of λ) $P^C_{z_A}(\uparrow|\lambda_1)$ simply by

$$P^C_{z_A}(\uparrow|\lambda_1) = \frac{n_{\uparrow,A,1}}{N_{A,1}}. \tag{5.8}$$

Bob determines (and Alice agrees)

$$P^C_{z_B}(\uparrow|\lambda_1) = \frac{n_{\uparrow,B,1}}{N_{B,1}} = \frac{n_{\uparrow,B,1}}{N_{A,1}}. \tag{5.9}$$

Note that these relations are only exact when N goes to infinity.

5.3.2 Different frames

The interesting situation arises when our couple chooses their axes z_A and z_B differently. The PHV approach then tells us the probability Bob measures is then perceived differently by Alice. Let us contemplate for a minute how the values of λ in the table above could change if Bob chooses his axes slightly different. If all Bob's values λ_1 and λ_2 change similarly to a unique same (different) value, say all λ_1s to λ_2s and vice versa, this does not make any real difference with the previous subsets: they would be merely inverted. This would not be different for the cases with more that two different λs. In a straightforward transformation as that, $\uparrow, \downarrow$ data would not be shuffled. Alice perception of Bob's probabilities to measure up or down would then never really differ from his own. In order for a difference to occur, we need the transformation $\Lambda(\lambda_1) = \lambda'_1$ to mix up the two subsets of Bob. This can only happen if only some λs change and some do not. We therefore suppose that there is a certain probability for a λ to take a different value in another frame. A possible new table of outcomes for the EPR Bohm set up could then be

trial	Alice measures along z_A		Bob measures along z_B	
	spin value	λ_A	spin value	λ_B
1	$\downarrow$	λ_1	$\uparrow$	λ_2
2	$\uparrow$	λ_2	$\uparrow$	λ_2
3	$\uparrow$	λ_2	$\downarrow$	λ_1
4	$\downarrow$	λ_1	$\downarrow$	λ_1
5	$\downarrow$	λ_1	$\uparrow$	λ_2
6	$\uparrow$	λ_1	$\uparrow$	λ_1
etc.				
N				

Given the measurement data, it is clear which values λ_1 of Bob Alice would interpret as λ_2: every time she measured λ_2 herself. Likewise, for the exact same trials, Bob would interpret Alice's λ_2 as λ_1. In that sense, the transformation from Bob to Alice is the exact inverse of the transformation from Alice to Bob, since the number of differently interpreted values is the same for Alice and Bob. However, we can of course not define the transformation on the basis of all these individual cases. We would like to have a general rule which tells Alice, without consulting her own outcomes of the λ or spin measurements, how to transform Bob's λ values.

Let us suppose it is purely a matter of probability whether a λ is perceived to have another value in another frame. It should then not matter which of the individual trials of Bob's λ_1 cases Alice would change to λ_2 cases as long as her adjustments would give rise to the same probabilities to measure the spin. So what we are interested in is the probability for Alice and Bob to measure λ_1 and λ_2 in a trial. They would determine it by, for large N

$$q_{A,1} = \frac{N_{A,1}}{N}, \qquad q_{A,2} = \frac{N_{A,2}}{N} = 1 - q_{A,1}, \qquad q_{B,1} = \frac{N_{B,1}}{N}, \qquad q_{B,2} = \frac{N_{B,2}}{N} = 1 - q_{B,1}, \tag{5.10}$$

where

$$N = N_{A,1} + N_{A,2} = N_{B,1} + N_{B,2}. \tag{5.11}$$

Respectively $q_{A,1}$ and $q_{B,1}$ then represent the probability for the source of shooting a λ_1 for Alice and Bob. We will refer to them as the source probabilities.

The way for Alice to transform Bob's data, if she would not have any λ measurements herself, would go in terms of different probabilities. We simply define the probability for a λ_1 in Bobs frame to become a λ_1 in Alice frame to be $P_{A,1}$. For it remain λ_1 is then automatically $1 - P_{A,1}$. Every time Bob measures λ_2, it has a probability $P_{A,2}$ to become a λ_1 and a probability $1 - P_{A,2}$ to remain a λ_2. In terms of measurement outcomes Alice could determine

$$P_{A,1} = \frac{M_{A,2}}{N_{B,1}} \qquad P_{A,2} = \frac{M_{A,1}}{N - N_{B,1}} \tag{5.12}$$

Where $M_{A,2}$ is determined as follows: Bob's subset defined by λ_1 is considered and from the corresponding λ values Alice measured all λ_2 cases are counted: this is $M_{A,2}$. Likewise, $M_{A,1}$ is obtained counting Alice's λ_1 values in the subset defined by Bob's λ_2. It can be seen to count that

$$N_{B,1} - M_{A,2} + M_{A,1} = N_{A,1} \qquad N - N_{B,1} - M_{A,1} + M_{A,2} = N - N_{A,1} \tag{5.13}$$

So Alice can transform Bob's data by changing the individual λ cases accordingly. Likewise Bob could do the same with Alice's data (if he would not know his λ data) using $P_{B,1}$ and $P_{B,2}$ (or $M_{B,1}$ and $M_{B,2}$). Summarized:

$$\Lambda_{AB}(\lambda_1) = \begin{cases} \lambda_2 & \text{with probability } P_{1,A} \\ \lambda_1 & \text{with probability } 1 - P_{1,A} \end{cases}$$

$$\Lambda_{AB}(\lambda_2) = \begin{cases} \lambda_1 & \text{with probability } P_{2,A} \\ \lambda_2 & \text{with probability } 1 - P_{2,A} \end{cases}$$

Bob could of course do the exact same thing taking as a basis his values,

$$\Lambda_{BA}(\lambda_1) = \begin{cases} \lambda_2 & \text{with probability } P_{1,B} \\ \lambda_1 & \text{with probability } 1 - P_{1,B} \end{cases}$$

$$\Lambda_{BA}(\lambda_2) = \begin{cases} \lambda_1 & \text{with probability } P_{2,B} \\ \lambda_2 & \text{with probability } 1 - P_{2,B} \end{cases}$$

So we literally need Alice (and Bob) to play a completely honest random number generator and make the adjustments accordingly, not looking at any spin outcomes. We can deduce that the size of the Bob's subset after Alice transforms it is, when the number of trials N is very large, expected to be

$$N_{B,1}^{A} = q_{A,1}N = N_{A,1}, \qquad N_{B,2}^{A} = q_{A,2}N = N_{A,2}. \tag{5.14}$$

Likewise Bob adjustment of Alice's subsets makes them of an expected size

$$N_{A,1}^{B} = q_{B,1}N = N_{B,1}, \qquad N_{A,2}^{B} = q_{B,2}N = N_{B,2}, \tag{5.15}$$

where it still counts that

$$N = N_{A,1}^{B} + N_{A,2}^{B} = N_{B,1}^{A} + N_{B,2}^{A}. \tag{5.16}$$

We would like investigate the effect of the transformation on the number of spin up and downs. In order to do that we introduce different variables from the Ps. It turns out to be more convenient in order to transform a number of λ_1 and λ_2 cases to use fractions p, as these can also take negative values. The interpretative relation between P and p we shall specify later. We introduce $p_{A,1}$ as the fraction of λ_1 cases Alice takes from $N_{B,1}$ and changes to λ_2. $1 - p_{A,1}$ then represents the fraction of $N_{B,1}$ she keeps as λ_1. Likewise she introduces the fraction $p_{A,2}$ as the fraction of λ_2 cases Alice takes from $N_{B,2}$ and changes to λ_1. If for example the fraction $p_{A,1}$ has a negative value, it simply means that Alice is not really changing any of the λ_1 cases to λ_2 cases: she is only doing the opposite: drawing λ_2s from the other subset and changing them to λ_1. Later calculations should then define the precise values of these ps as functions of the angle between Alice and Bob and other quantities.

We would expect that Alice, in her arbitrary adjustments, changes the same fraction $p_{1,A}$ of λ_1 spin ups to λ_2 spin ups as she would change λ_1 spin downs to λ_2 spin downs, especially if the number of adjustments becomes very large. So if $n_{\uparrow,1,B}$ is the number of Bob's spin ups in his subset defined by λ_1 and $n_{\uparrow,2,B}$ is the number of Bob's spin ups in his subset defined by λ_2, writing these in a two dimensional vector, we have the transformation rule

$$\begin{pmatrix} n^A_{\uparrow,B,1} \\ n^A_{\uparrow,B,2} \end{pmatrix} = \begin{pmatrix} 1-p_{1,A} & p_{2,A} \\ p_{1,A} & 1-p_{2,A} \end{pmatrix} \begin{pmatrix} n_{\uparrow,B,1} \\ n_{\uparrow,B,2} \end{pmatrix}$$

Where $n^A_{\uparrow,B,1}, n^A_{\uparrow,B,2}$ are then the number of spin ups in Bob's subsets defined by λ_1 and λ_2 as Alice perceives them. The transformation is necessarily similar for a vector with $n^B_{\downarrow,1,B}$ and $n^B_{\downarrow,2,B}$. The same should count for Bob with his fractions $p_{1,B}$ and $p_{2,B}$

$$\begin{pmatrix} n^B_{\uparrow,A,1} \\ n^B_{\uparrow,A,2} \end{pmatrix} = \begin{pmatrix} 1-p_{1,B} & p_{2,B} \\ p_{1,B} & 1-p_{2,B} \end{pmatrix} \begin{pmatrix} n_{\uparrow,A,1} \\ n_{\uparrow,A,2} \end{pmatrix}$$

This transformation matrix from Bob to Alice should then be the inverse of the transformation from Alice to Bob. The inverse of the latter is given by

$$\frac{1}{p_{1,A}+p_{2,A}-1} \begin{pmatrix} p_{2,A}-1 & p_{2,A} \\ p_{1,A} & p_{1,A}-1 \end{pmatrix}$$

So we make the identification

$$1-p_{1,B} = \frac{p_{2,A}-1}{p_{1,A}+p_{2,A}-1} = \frac{1}{\frac{p_{1,A}}{p_{2,A}-1}+1}, \qquad 1-p_{2,B} = \frac{1}{\frac{p_{2,A}}{p_{1,A}-1}+1}, \tag{5.17}$$

$$p_{1,B} = \frac{1}{\frac{p_{2,A}-1}{p_{1,A}}+1}, \qquad p_{2,B} = \frac{1}{\frac{p_{1,A}-1}{p_{2,A}}+1}, \tag{5.18}$$

which fixing the values of $p_{B,1}$ and $p_{B,2}$ in terms of $p_{A,1}$ and $p_{A,2}$. In the following we will therefore only use the variables $p_{A,1} = p_1$ and $p_{A,2} = p_2$.

What we are really interested in are the conditional probabilities to measure spin up or spin down given a certain value of λ and a direct transformation of those. We know that $P^C_{z_B}(\uparrow|\lambda_1)$ is proportional to the number $n_{\uparrow,B,1}$. Alice would determine this probability on the basis of Bob's transformed λ data. The conditional probability for Bob to measure spin up as perceived by Alice $P^C_{z_B}(\uparrow|\lambda_1)^A$ is similarly determined from the new number $n^A_{\uparrow,B,1}$ of $\uparrow$'s in the new subset defined by λ_1. So let us define, for N going to infinity:

$P^C_{z_B}(\uparrow|\lambda_1) = n_{\uparrow,B,1}/N_{B,1}$ to be the conditional probability for Bob to measure spin up as determined by him in his frame, given a value of λ_1.

$P^C_{z_B}(\uparrow|\lambda_1)^A = n^A_{\uparrow,B,1}/N^A_{B,1}$ to be the conditional probability for Bob to measure up, as determined by Alice in her frame, given a value λ_1.

$P^C_{z_B}(\uparrow|\lambda_1) = n_{\uparrow,A,1}/N_{A,1}$ to be the conditional probability for Alice to measure spin up as determined by her in his frame, given a value λ_1.

$P^C_{z_A}(\uparrow|\lambda_1)^B = n^B_{\uparrow,A,1}/N^B_{A,1}$ to be the conditional probability for Alice to measure up, as determined by Bob in his frame, given a value λ_1.

The definitions are similar for λ_2. Note that Alice and Bob do not agree about the cases where λ_1 is given, therefore, they do not agree which conditional probability applied to which case. We already know how to calculate

$$n^A_{\uparrow,B,1} = (1-p_1)n_{\uparrow,B,1} + p_2 n_{\uparrow,B,2}. \tag{5.19}$$

Likewise $N^A_{B,1}$ is the size of Bob's subset defined by $\lambda_{1,B}$ as Alice perceives it

$$\begin{aligned} N^A_{B,1} &= n^A_{\uparrow,B,1} + n^A_{\downarrow,B,1} \\ &= (1-p_1)n_{\uparrow,B,1} + p_2 n_{\uparrow,B,2} + (1-p_1)n_{\downarrow,B,1} + p_2 n_{\downarrow,B,2} \\ &= (1-p_1)N_{B,1} + p_2(N - N_{B,1}), \end{aligned} \tag{5.20}$$

$$P^C_{z_B}(\uparrow|\lambda_1)^A = (1-p_1)P^C_{z_B}(\uparrow|\lambda_1)\frac{N_{B,1}}{N^A_{B,1}} + p_2 P^C_{z_B}(\uparrow|\lambda_2)\frac{N-N_{B,1}}{N^A_{B,1}}. \tag{5.21}$$

So that

$$P^C_{z_B}(\uparrow|\lambda_1)^A = \frac{(1-p_1)P^C_{z_B}(\uparrow|\lambda_1)N_{B,1} + p_2 P^C_{z_B}(\uparrow|\lambda_2)(N-N_{B,1})}{(1-p_1)N_{B,1} + p_2(N-N_{B,1})}, \tag{5.22}$$

$$P^C_{z_B}(\uparrow|\lambda_2)^A = \frac{(1-p_2)P^C_{z_B}(\uparrow|\lambda_2)(N-N_{B,1}) + p_1 P^C_{z_B}(\uparrow|\lambda_1)N_1}{N - (1-p_1)N_{B,1} - p_2(N-N_{B,1})}. \tag{5.23}$$

As a quick way to see how Alice's probabilities transform, we simply make the substitutions from Alice's fractions to Bob's

$$1-p_1 \to \frac{p_2-1}{p_1+p_2-1} = \frac{1}{\frac{p_1}{p_2-1}+1}, \qquad 1-p_2 \to= \frac{1}{\frac{p_2}{p_1-1}+1}, \tag{5.24}$$

$$p_1 \to \frac{1}{\frac{p_2-1}{p_1}+1} \qquad p_2 \to= \frac{1}{\frac{p_1-1}{p_2}+1} \tag{5.25}$$

And we take $N_{A,1}$, the size of her subset defined by λ_1 as Alice measures it, to play the role of $N_{B,1}$. So we obtain for a transformation of Alice's probabilities:

$$P^C_{z_A}(\uparrow|\lambda_1)^B = \frac{\frac{N_{A,1}}{\frac{p_1}{p_2-1}+1}P^C_{z_A}(\uparrow|\lambda_1)+\frac{N-N_{A,1}}{\frac{p_1-1}{p_2}+1}P^C_{z_A}(\uparrow|\lambda_2)}{\frac{N_{A,1}}{\frac{p_1}{p_2-1}+1}+\frac{N-N_{A,1}}{\frac{p_1-1}{p_2}+1}} \tag{5.26}$$

$$P^C_{z_A}(\uparrow|\lambda_1)^B = \frac{N_{A,1}(p_2-1)P^C_{z_A}(\uparrow|\lambda_1)+(N-N_{A,1})p_2P^C_{z_A}(\uparrow|\lambda_2)}{N_{A,1}(p_2-1)+(N-N_{A,1})p_2} \tag{5.27}$$

$$P^C_{z_A}(\uparrow|\lambda_1^B) = \frac{N_{A,1}(p_2-1)P^C_{z_A}(\uparrow|\lambda_1)+(N-N_{A,1})p_2P^C_{z_A}(\uparrow|\lambda_2)}{Np_2-N_{A,1}} \tag{5.28}$$

And for her $P^C_{z_A}(\uparrow|\lambda_2)$

$$P^C_{z_A}(\uparrow|\lambda_2)^B = \frac{(p_1-1)P^C_{z_A}(\uparrow|\lambda_2)(N-N_{A,1})+p_1P^C_{z_A}(\uparrow|\lambda_1)N_{A,1}}{(p_1-1)N+N_{A,1}} \tag{5.29}$$

So we see that we could define

$$N^B_{A,1} = \frac{p_2N-N_{A,1}}{p_1+p_2-1} \qquad N-N^B_{A,1} = \frac{(p_1-1)N+N_{A,1}}{p_1+p_2-1} \tag{5.30}$$

Note again that these results are really expectation values, which only become exact when the number of trials N is very large. We observe for p_1 and p_2 both equal to 0, there is no transformation at all, which of course desired for $z_A = z_B$, when Alice and Bob are in the same frame. We also observe that if $P^C_{z_B}(\uparrow|\lambda_1) = P^C_{z_B}(\uparrow|\lambda_2)$, no matter what transformation is applied, $P^C_{z_B}(\uparrow|\lambda_1)^A = P^C_{z_B}(\uparrow|\lambda_1)$. This makes sense, because if λ_1 should really be the same as λ_2 if it predicts the same probability for measuring $\uparrow$.

In the next subsections it will proven useful to write the transformation of the probabilities too as a matrix multiplication (we will stop writing the subscript C for conditional probabilities, unless there is confusion possible):

$$\begin{pmatrix} P_{z_B}(\uparrow|\lambda_1)^A \\ P_{z_B}(\uparrow|\lambda_2)^A \end{pmatrix} = \begin{pmatrix} (1-p_1)\frac{N_{B,1}}{N^A_{B,1}} & p_2\frac{N-N_{B,1}}{N^A_{B,1}} \\ p_1\frac{N_{B,1}}{N-N^A_{B,1}} & (1-p_2)\frac{N-N_{B,1}}{N-N^A_{B,1}} \end{pmatrix} \begin{pmatrix} P_{z_B}(\uparrow|\lambda_1) \\ P_{z_B}(\uparrow|\lambda_2) \end{pmatrix},$$

where $N^A_{B,1} = (1-p_1)N_{B,1} + p_2(N-N_{B,1})$. We will call this probability vector $\vec{v}_{z_B,\uparrow}$ and the transformation matrix M_{BA}, for spin down it is $\vec{v}_{z_B,\downarrow}$. $\vec{v}_{z_B,\downarrow}$ transforms with the same matrix M_{BA}. Alice's transformation is given by

$$\begin{pmatrix} P_{z_A}(\uparrow|\lambda_1^B) \\ P_{z_A}(\uparrow|\lambda_2^B) \end{pmatrix} = \begin{pmatrix} (p_2-1)\frac{N_{A,1}}{p_2N-N_{A,1}} & p_2\frac{N-N_{A,1}}{p_2N-N_{A,1}} \\ p_1\frac{N_{A,1}}{(p_1-1)N+N_{A,1}} & (p_1-1)\frac{N-N_1}{(p_1-1)N+N_{A,1}} \end{pmatrix} \begin{pmatrix} P_{z_A}(\uparrow|\lambda_1) \\ P_{z_A}(\uparrow|\lambda_2) \end{pmatrix}$$

or

$$\begin{pmatrix} P_{z_A}(\uparrow|\lambda_1^B) \\ P_{z_A}(\uparrow|\lambda_2^B) \end{pmatrix} = \frac{1}{p_1+p_2-1} \begin{pmatrix} (p_2-1)\frac{N_{A,1}}{N_{A,1}^B} & p_2\frac{N-N_{A,1}}{N_{A,1}^B} \\ p_1\frac{N_{A,1}}{N-N_{A,1}^B} & (p_1-1)\frac{N-N_1}{N-N_{A,1}^B} \end{pmatrix} \begin{pmatrix} P_{z_A}(\uparrow|\lambda_1) \\ P_{z_A}(\uparrow|\lambda_2) \end{pmatrix}.$$

For vectors $\vec{v}_{z_A,\uparrow}$, $\vec{v}_{z_A,\downarrow}$. Note that the transformation matrix M_{AB} for Alice's probabilities is equal to M_{BA}^{-1} if the following conditions are satisfied:

$$N_{B,1} = N_{A,1}^B \qquad N_{A,1} = N_{B,1}^A \tag{5.31}$$

(Which can be shown to be equivalent after substitutions of (5.20) and (5.30).) We see that this is in perfect accordance with the relations (5.14) and (5.15).

5.4 Consistency with probabilities from QM

If the PHV model can be taken as a serious alternative to QM, it should at least be able to reproduce the probabilities which QM would predict. To be able to do this, a couple restrictions are put on the PHV. For simplicity we stay with the possibility where Alice and Bob measure the same λ on their meters and where there is only two possible values λ_1 and λ_2 into play. The source produces λ_1 in the limit where N goes to infinity, from Alice point of view with probability $q_{A,1} = N_{A,1}/N$, from Bob's point of view, this is $q_{B,1} = N_{B,1}/N$. We first check whether agreement between Alice and Bob can be reached at all about probabilities if the supposed knowledge about λ_1 and λ_2 is ignored. This is done first for the marginal and then for the joint probabilities. It will fix the form of the transformation of λ uniquely in terms of $q_{1,B}$. In the next section it is investigated if this PHV model could reproduce the probabilities which QM would predict for the EPR Bohm set up.

5.4.1 Checking marginal probabilities

It follows straightforwardly that the marginal probability for Alice to measure spin up should satisfy, according to her

$$P_{z_A}^{M,QM}(\uparrow) = P_{z_A}^C(\uparrow|\lambda_{A,1})\frac{N_{A,1}}{N} + P_{z_A}^C(\uparrow|\lambda_{A,2})\frac{N-N_{A,1}}{N}. \tag{5.32}$$

She simply takes the probability to have λ_1 and multiplies it with the conditional probability for her to measure up given a value λ_1. This she adds to the probability obtained by the same procedure for λ_2, as in a QM calculation, no distinction is made between λ_1 and λ_2. The marginal probability for Bob to measure up, according to Alice, is similarly given by (using that $N^A_{B,1} = (1-p_1)N_{B,1} + p_2(N - N_{B,1})$):

$$\begin{aligned} P^{M,QM}_{z_B}(\uparrow) &= P^C_{z_B}(\uparrow|\lambda_1)^A \frac{N^A_{B,1}}{N} + P^C_{z_B}(\uparrow|\lambda_2)^A \frac{N - N^A_{B,1}}{N} \\ &= \frac{(1-p_1)P^C_{z_B}(\uparrow|\lambda_1)N_{B,1} + p_2 P^C_{z_B}(\uparrow|\lambda_2)(N - N_{B,1})}{N} + \\ &\quad \frac{(1-p_2)P^C_{z_B}(\uparrow|\lambda_2)(N - N_{B,1}) + p_1 P^C_{z_B}(\uparrow|\lambda_1)N_{B,1}}{N} \\ &= P^C_{z_B}(\uparrow|\lambda_1)\frac{N_{B,1}}{N} + P^C_{z_B}(\uparrow|\lambda_2)\frac{N - N_{B,1}}{N}. \end{aligned} \tag{5.33}$$

So this automatically agrees with the value as calculated by Bob from his frame:

$$P^{M,QM}_{z_B}(\uparrow) = P^C_{z_B}(\uparrow|\lambda_{B,1})\frac{N_{B,1}}{N} + P^C_{z_B}(\uparrow|\lambda_{B,2})\frac{N - N_{B,1}}{N} \tag{5.34}$$

To check Alice's marginal probability as seen by Bob, we note this time we take

$$N^B_{A,1} = \frac{Np_2 - N_{A,1}}{p_2 + p_1 - 1} \qquad N - N^B_{A,1} = \frac{N(p_1 - 1) + N_{A,1}}{p_2 + p_1 - 1} \tag{5.35}$$

Alice's marginal probability to measure spin up is then given, from Bob's frame, by

$$\begin{aligned} P^{M,QM}_{z_A}(\uparrow) &= P^C_{z_A}(\uparrow|\lambda_1)^B \frac{N^B_{A,1}}{N} + P^C_{z_A}(\uparrow|\lambda_2)^B \frac{N - N^B_{A,1}}{N} \\ &= \frac{(p_2-1)P^C_{z_A}(\uparrow|\lambda_{A,1})N_{A,1} + p_2 P^C_{z_A}(\uparrow|\lambda_{A,2})(N - N_{A,1})}{N(p_2 + p_1 - 1)} + \\ &\quad \frac{(p_1-1)P^C_{z_A}(\uparrow|\lambda_{A,2})(N - N_{A,1}) + p_1 P^C_{z_A}(\uparrow|\lambda_{A,1})N_{A,1}}{N(p_2 + p_1 - 1)} \\ &= P^C_{z_A}(\uparrow|\lambda_{A,1})\frac{N_{A,1}}{N} + P^C_{z_A}(\uparrow|\lambda_{A,2})\frac{N - N_{A,1}}{N}, \end{aligned} \tag{5.36}$$

so for her marginal probability too there is automatic agreement. The marginal probabilities for measuring spin down follow similarly.

5.4.2 Checking joint probabilities

At first there seems to be a problem if we would want to check the joint probabilities in terms of the hidden variable λ: if Alice and Bob choose their axes different, there is no agreement about which λ was the case at which trial. Yet for Alice's and Bob's approaches separately, the joint probability can be calculated. In the end these two methods should result in the same (QM) outcome.

Every time Alice measures a value λ_1, she knows with certainty Bob would have measured λ_1 if he were in her frame. If for the same trial Bob measures a λ_2 in another frame, the transformation she will apply amounts to changing it to λ_1. The λ is then rearranged to the situation when Alice and Bob were both in Alice frame. For this situation, it is easy for Alice to find the probability for them both to measure spin up given a value λ_1: she multiplies the conditional probability for her to measure up given λ_1 times the conditional probability for Bob to measure up given λ_1 times the probability to have the value λ_1 as perceived in Alice frame. So indeed the probabilities factorize as should be the case if the correlation is classical. The joint probability as QM predicts is then again obtained if we are indifferent to the difference between the λ_1 and λ_2 cases. The joint probabilities for the two cases are simply summed. So the joint probability for both Alice and Bob to measure spin up according to QM is then given by, from Alice's frame, for N going to infinity

$$P^{J,QM}(\uparrow,\uparrow) = P^C_{z_A}(\uparrow|\lambda_1)P^C_{z_B}(\uparrow|\lambda_1)^A\frac{N_{A,1}}{N} + P^C_{z_A}(\uparrow|\lambda_2)P^C_{z_B}(\uparrow|\lambda_2)^A\frac{N-N_{A,1}}{N}. \quad (5.37)$$

We stop writing the 'Alice' and 'Bob' labels for the λs, they are assumed to be the same when Alice and Bob are in the same frame. Using the results from the previous section,

$$\begin{aligned} P^{J,QM}(\uparrow,\uparrow) &= P^C_{z_A}(\uparrow|\lambda_1)\frac{(1-p_1)P^C_{z_B}(\uparrow|\lambda_1)N_{B,1} + p_2P^C_{z_B}(\uparrow|\lambda_2)(N-N_{B,1})}{(1-p_1)N_{B,1}+p_2(N-N_{B,1})}\frac{N_{A,1}}{N} \\ &+ P^C_{z_A}(\uparrow|\lambda_2)\frac{(1-p_2)P^C_{z_B}(\uparrow|\lambda_1)(N-N_{B,1}) + p_1P^C_{z_B}(\uparrow|\lambda_2)N_{B,1}}{N-(1-p_1)N_{B,1}-p_2(N-N_{B,1})}\frac{N-N_{A,1}}{N}. \quad (5.38) \end{aligned}$$

Bob would claim the probability is given by (again, for N going to infinity))

$$P^{J,QM}(\uparrow,\uparrow) = P^C_{z_A}(\uparrow|\lambda_1)^BP^C_{z_B}(\uparrow|\lambda_1)\frac{N_{B,1}}{N} + P^C_{z_A}(\uparrow|\lambda_2)^BP^C_{z_B}(\uparrow|\lambda_2)\frac{N-N_{B,1}}{N}, \quad (5.39)$$

which is written

$$\begin{aligned} P^{J,QM}(\uparrow,\uparrow) = P^C_{z_B}(\uparrow|\lambda_1)\frac{N_{A,1}(p_2-1)P^C_{z_A}(\uparrow|\lambda_1)+(N-N_{A,1})p_2P^C_{z_A}(\uparrow|\lambda_2)}{Np_2-N_{A,1}}\frac{N_{B,1}}{N} \\ +P^C_{z_B}(\uparrow|\lambda_2)\frac{(p_1-1)P^C_{z_A}(\uparrow|\lambda_2)(N-N_{A,1})+p_1P^C_{z_A}(\uparrow|\lambda_1)N_{A,1}}{(p_1-1)N+N_{A,1}}\frac{N-N_{B,1}}{N}. \quad (5.40)\end{aligned}$$

Alice and Bob had better agree on their QM, so we require (5.38)=(5.40) in the limit where N goes to infinity. If the relation (5.31) is satisfied, again using (5.35) this simplifies to

$$\begin{aligned} P^C_{z_A}(\uparrow|\lambda_1)\big((1-p_1)P^C_{z_B}(\uparrow|\lambda_1)N_{B,1}+p_2P^C_{z_B}(\uparrow|\lambda_2)(N-N_{B,1})\big) &+ \\ P^C_{z_A}(\uparrow|\lambda_2)((1-p_2)P^C_{z_B}(\uparrow|\lambda_1)(N-N_{B,1})+p_1P^C_{z_B}(\uparrow|\lambda_2)N_{B,1}) &= \\ \frac{1}{p_1+p_2-1}\big(P^C_{z_B}(\uparrow|\lambda_1)(N_{A,1}(p_2-1)P^C_{z_A}(\uparrow|\lambda_1)+(N-N_{A,1})p_2P^C_{z_A}(\uparrow|\lambda_2)) &+ \\ P^C_{z_B}(\uparrow|\lambda_2)((p_1-1)P^C_{z_A}(\uparrow|\lambda_2)(N-N_{A,1})+p_1P^C_{z_A}(\uparrow|\lambda_1)N_{A,1})\big) & \quad (5.41)\end{aligned}$$

$$\begin{aligned} P^C_{z_A}(\uparrow|\lambda_1)P^C_{z_B}(\uparrow|\lambda_1)\big((1-p_1)N_{B,1}-\frac{N_{A,1}(p_2-1)}{p_1+p_2-1}\big) &+ \\ P^C_{z_B}(\uparrow|\lambda_2)P^C_{z_A}(\uparrow|\lambda_1)\big(p_2(N-N_{B,1})-\frac{p_1N_{A,1}}{p_1+p_2-1}\big) &+ \\ P^C_{z_A}(\uparrow|\lambda_2)P^C_{z_B}(\uparrow|\lambda_1)\big((1-p_2)(N-N_{B,1})-\frac{(N-N_{A,1})p_2}{p_1+p_2-1}\big) &+ \\ P^C_{z_B}(\uparrow|\lambda_2)P^C_{z_A}(\uparrow|\lambda_2)\big(p_1N_{B,1}-\frac{(N-N_{A,1})(p_1-1)}{p_1+p_2-1}\big) &= 0 \quad (5.42)\end{aligned}$$

Before we start simplifying this, we note that all the other probabilities just consist of replacing $P^C_{z_B}(\uparrow|\lambda_2)$ for $P^C_{z_B}(\downarrow|\lambda_2)$ or the other way around: $P^C_{z_A}(\uparrow|\lambda_2)$ for $P^C_{z_A}(\downarrow|\lambda_2)$. For the joint probability for both to measure down both substitutions are made. One ends up with three more similar equations to (5.42). Another important aspect of this equation is that it should be satisfied for any value of the P^Cs, as there should be nothing special about the orientations z_A and z_B (on which these conditional probabilities may depend). We therefore require, and thereby satisfying automatically the expressions for the other joint probabilities as well, that

$$\begin{aligned}(1-p_1)N_{B,1}-\frac{N_{A,1}(p_2-1)}{p_1+p_2-1} &= 0,\\ (p_2(N-N_{B,1})-\frac{p_1N_{A,1}}{p_1+p_2-1}) &= 0,\\ (1-p_2)(N-N_{B,1})-\frac{(N-N_{A,1})p_2}{p_1+p_2-1} &= 0,\\ p_1N_{B,1}-\frac{(N-N_{A,1})(p_1-1)}{p_1+p_2-1} &= 0. \end{aligned} \tag{5.43}$$

These are actually only two equations which reduce to the following requirements (using again $N_{A,1}=(1-p_1)N_{B,1}+p_2(N-N_{B,1})$)

$$\frac{p_1(1-p_1)}{p_2(1-p_2)}=1-\frac{N}{N_{B,1}}, \qquad \frac{p_1p_2}{p_1+p_2-1}=1-\frac{N_{B,1}}{N}. \tag{5.44}$$

These two equations for two unknowns p_1 and p_2 can be solved completely. We find

$$p_1=2(1-\frac{N_{B,1}}{N}) \qquad p_2=1-2\frac{N_{B,1}}{N}. \tag{5.45}$$

Note that there is really only dependence on the ratio $N_{B,1}/N$, which is desirable, because now the limit $N\to\infty$ can be taken easily, obtaining.

$$p_1=2(1-q_{B,1}) \qquad p_2=1-2q_{B,1} \tag{5.46}$$

Note that $p_1-1=p_2$.

So we conclude that with these values for p_1 and p_2, for a given $q_{B,1}$ (which probably depends on z_A,z_B) consistency is obtained with QM. Note that it is not yet checked if the predictions of QM can be made to coincide exactly (with the right angle dependence). We will investigate this in the next subsection. Finally we derive a relation between the two different source probabilities $q_{B,1}$ and $q_{A,1}$ which we will need for future reference. Setting the relations in (5.35) equal to respectively $N_{B,1}$ and $N-N_{B,1}$ (which should count according to (5.31)) and taking the large N limit

$$q_{B,1}=\frac{p_2-q_{A,1}}{p_2+p_1-1}, \qquad 1-q_{B,1}=\frac{p_1-1+q_{A,1}}{p_2+p_1-1}. \tag{5.47}$$

Dividing the right equation by the left one, we obtain

$$\frac{1}{q_{B,1}}-1=\frac{p_1-1+q_{A,1}}{p_2-q_{A,1}}, \tag{5.48}$$

$$p_2(1 - q_{B,1}) + q_{B,1}(1 - p_1) = q_{A,1}. \tag{5.49}$$

Using the results from (5.45), this further reduces to

$$q_{A,1} = (2q_{B,1} - 1)^2. \tag{5.50}$$

Note that for $q_{B,1} = 1/4$, it happens to be exactly equal to $q_{A,1}$. However, the reverse statement does not count: we know indeed that the source probabilities are equal to Bob and Alice when they are in the same frame, but then it also counts that p_1 and p_2 are 0, which would be a conflicting statement with (5.46) (and even more for $q = 1/4$). It seems that the case of different frames is fundamentally different from the case of similar frames. It would be nice if a somewhat 'smooth transition' between the two can be identified. This will not be investigated here.

5.4.3 Producing the Quantum Mechanical angle dependence

That there is agreement between Alice and Bob about the marginal and joint probabilities in the previous subsection was only a necessary requirement to make the PHV consistent with QM . We will now investigate if these probabilities can also be given the right form, the right form being the one predicted by QM. We will keep on working within the EPR Bohm set up and we will start consider a general two particle bipartite state. In the next section we will try to obtain consistency for some particular examples.

Suppose we have a state

$$|\Psi\rangle = a\,|z\uparrow\rangle\,|z\uparrow\rangle + b\,|z\uparrow\rangle\,|z\downarrow\rangle + c\,|z\downarrow\rangle\,|z\uparrow\rangle + d\,|z\downarrow\rangle\,|z\downarrow\rangle \tag{5.51}$$

Alice and Bob are of course free to choose their axes different from this z. In the first appendix it is shown how the state can then be rewritten in terms of the angles these axes make with this original z axes, so that it can be conveniently calculated how the probabilities depend on these angles. Let us suppose Alice chooses her axis instead $\vec{n}(\theta_1, \phi_1)$, along Bob chooses his axis along $\vec{m}(\theta_2, \phi_2)$, so the state becomes in those terms

$$\begin{aligned} |\Psi\rangle = \alpha(\theta_1,\phi_1,\theta_2,\phi_2)\,|\vec{n}\uparrow\rangle_A\,|\vec{m}\uparrow\rangle_B + \beta(\theta_1,\phi_1,\theta_2,\phi_2)\,|\vec{n}\uparrow\rangle_A\,|\vec{m}\downarrow\rangle_B +\\ \gamma(\theta_1,\phi_1,\theta_2,\phi_2)\,|\vec{n}\downarrow\rangle_A\,|\vec{m}\uparrow\rangle_B + \delta(\theta_1,\phi_1,\theta_2,\phi_2)\,|\vec{n}\downarrow\rangle_A\,|\vec{m}\downarrow\rangle_B \end{aligned} \tag{5.52}$$

Where the coefficients are given by

$$\begin{aligned}
\alpha(\theta_1,\phi_1,\theta_2,\phi_2) &= e^{\frac{i(\phi_1+\phi_2)}{2}}cos\frac{\theta_1}{2}cos\frac{\theta_2}{2}a + e^{\frac{i(\phi_1-\phi_2)}{2}}cos\frac{\theta_1}{2}sin\frac{\theta_2}{2}b \\
&\quad +e^{\frac{i(\phi_2-\phi_1)}{2}}sin\frac{\theta_1}{2}cos\frac{\theta_2}{2}c + e^{-\frac{i(\phi_1+\phi_2)}{2}}sin\frac{\theta_1}{2}sin\frac{\theta_2}{2}d, \quad (5.53)\\
\beta(\theta_1,\phi_1,\theta_2,\phi_2) &= -e^{\frac{i(\phi_1+\phi_2)}{2}}cos\frac{\theta_1}{2}sin\frac{\theta_2}{2}a + e^{\frac{i(\phi_1-\phi_2)}{2}}cos\frac{\theta_1}{2}cos\frac{\theta_2}{2}b \\
&\quad -e^{\frac{i(\phi_2-\phi_1)}{2}}sin\frac{\theta_1}{2}sin\frac{\theta_2}{2}c + e^{-\frac{i(\phi_1+\phi_2)}{2}}sin\frac{\theta_1}{2}cos\frac{\theta_2}{2}d, \quad (5.54)\\
\gamma(\theta_1,\phi_1,\theta_2,\phi_2) &= -e^{\frac{i(\phi_1+\phi_2)}{2}}sin\frac{\theta_1}{2}cos\frac{\theta_2}{2}a - e^{\frac{i(\phi_1-\phi_2)}{2}}sin\frac{\theta_1}{2}sin\frac{\theta_2}{2}b \\
&\quad +e^{\frac{i(\phi_2-\phi_1)}{2}}cos\frac{\theta_1}{2}cos\frac{\theta_2}{2}c + e^{-\frac{i(\phi_1+\phi_2)}{2}}cos\frac{\theta_1}{2}sin\frac{\theta_2}{2}d, \quad (5.55)\\
\delta(\theta_1,\phi_1,\theta_2,\phi_2) &= e^{\frac{i(\phi_1+\phi_2)}{2}}sin\frac{\theta_1}{2}sin\frac{\theta_2}{2}a - e^{\frac{i(\phi_1-\phi_2)}{2}}sin\frac{\theta_1}{2}cos\frac{\theta_2}{2}b \\
&\quad -e^{\frac{i(\phi_2-\phi_1)}{2}}cos\frac{\theta_1}{2}sin\frac{\theta_2}{2}c + e^{-\frac{i(\phi_1+\phi_2)}{2}}cos\frac{\theta_1}{2}cos\frac{\theta_2}{2}d. \quad (5.56)
\end{aligned}$$

For economy of writing we will no longer indicate the angle dependence of the coefficients. We are still considering how QM then tells us how the probabilities are calculated given such a state. Both measure the different QM joint probabilities

$$\begin{aligned}
P^{QM}_{z_B,z_A}(\uparrow,\uparrow) &= |\alpha|^2,\\
P^{QM}_{z_B,z_A}(\uparrow,\downarrow) &= |\beta|^2,\\
P^{QM}_{z_B,z_A}(\downarrow,\uparrow) &= |\gamma|^2,\\
P^{QM}_{z_B,z_A}(\downarrow,\downarrow) &= |\delta|^2. \quad (5.57)
\end{aligned}$$

As is clear from these joint probabilities, the marginal probabilities for respectively Alice and Bob to measure spin up is given by

$$P^{M,QM}_{z_A}(\uparrow) = |\alpha|^2 + |\beta|^2, \qquad P^{M,QM}_{z_B} \uparrow) = |\alpha|^2 + |\gamma|^2, \quad (5.58)$$

and for respectively Alice and Bob to measure spin down

$$P^{M,QM}_{z_A}(\downarrow) = |\delta|^2 + |\gamma|^2, \qquad P^{M,QM}_{z_B}(\downarrow) = |\delta|^2 + |\beta|^2. \quad (5.59)$$

We will now simply take the expression from the previous subsection and set them equal to these probabilities, where equality counts again in the large N limit. We saw previously that the marginal probability for Alice to measure up, both for Alice and Bob, is given by

$$P^{M,QM}_{z_A}(\uparrow) = P^C_{z_A}(\uparrow|\lambda_1)\frac{N_{A,1}}{N} + P^C_{z_A}(\uparrow|\lambda_2)\frac{N-N_{A,1}}{N}. \tag{5.60}$$

Using (5.30) and (5.45),

$$P^{M,QM}_{z_B}(\uparrow) = P^C_{z_A}(\uparrow|\lambda_1)(4\frac{N^2_{B,1}}{N^2} - 4\frac{N_{B,1}}{N} + 1) + P^C_{z_A}(\uparrow|\lambda_2)(4\frac{N_{B,1}}{N} - 4\frac{N^2_{B,1}}{N^2}). \tag{5.61}$$

We now impose this should be equal to $|\alpha|^2 + |\beta|^2$, (in what follows we will not write the angle dependence) so then we derive

$$P^C_{z_A}(\uparrow|\lambda_2) = \frac{P^C_{z_A}(\uparrow|\lambda_1)(\frac{N^2_{B,1}}{N^2} - \frac{N_{B,1}}{N} + \frac{1}{4}) - \frac{|\alpha|^2+|\beta|^2}{4}}{\frac{N^2_{B,1}}{N^2} - \frac{N_{B,1}}{N}}. \tag{5.62}$$

As Alice's marginal probability to measure spin down becomes likewise

$$P^C_{z_A}(\downarrow|\lambda_2) = \frac{P^C_{z_A}(\downarrow|\lambda_1)(\frac{N^2_{B,1}}{N^2} - \frac{N_{B,1}}{N} + \frac{1}{4}) - \frac{|\delta|^2+|\gamma|^2}{4}}{\frac{N^2_{B,1}}{N^2} - \frac{N_{B,1}}{N}}. \tag{5.63}$$

Bob's marginal probability to measure spin up is given by

$$P^{M,QM}_{z_B}(\uparrow) = P^C_{z_B}(\uparrow|\lambda_{B,1})\frac{N_{B,1}}{N} + P^C_{z_B}(\uparrow|\lambda_{B,2})\frac{N-N_{B,1}}{N}, \tag{5.64}$$

which should also be $|\alpha|^2 + |\gamma|^2$, therefore

$$P^C_{z_B}(\uparrow|\lambda_2) = \frac{P^C_{z_B}(\uparrow|\lambda_1)N_{B,1}/N - (|\alpha|^2+|\gamma|^2)}{N_{B,1}/N - 1}, \tag{5.65}$$

$$P^C_{z_B}(\downarrow|\lambda_2) = \frac{P^C_{z_B}(\downarrow|\lambda_1)N_{B,1}/N - (|\delta|^2+|\beta|^2)}{N_{B,1}/N - 1}. \tag{5.66}$$

We have seen in the previous subsection that the joint probability for both to measure up according to the PHV model is given, as Bob calculates it from his frame

$$\begin{aligned} P^{QM}_{z_B,z_A}(\uparrow,\uparrow) = & \quad P^C_{z_B}(\uparrow|\lambda_1)\frac{N_{A,1}(p_2-1)P^C_{z_A}(\uparrow|\lambda_1) + (N-N_{A,1})p_2P^C_{z_A}(\uparrow|\lambda_2)}{Np_2 - N_{A,1}}\frac{N_{B,1}}{N} \\ + & \quad P^C_{z_B}(\uparrow|\lambda_2)\frac{(p_1-1)P^C_{z_A}(\uparrow|\lambda_2)(N-N_{A,1}) + p_1P^C_{z_A}(\uparrow|\lambda_1)N_{A,1}}{(p_1-1)N + N_{A,1}}\frac{N-N_{B,1}}{N}. \end{aligned} \tag{5.67}$$

Substituting what we know from (5.30) and (5.31),

$$\begin{aligned} P^{QM}_{z_B,z_A}(\uparrow,\uparrow) &= \frac{1}{N(p_2+p_1-1)} \\ &\big(P^C_{z_B}(\uparrow|\lambda_1)(N_{A,1}(p_2-1)P^C_{z_A}(\uparrow|\lambda_1)+(N-N_{A,1})p_2P^C_{z_A}(\uparrow|\lambda_2)) \\ &+ P^C_{z_B}(\uparrow|\lambda_2)((p_1-1)P^C_{z_A}(\uparrow|\lambda_2)(N-N_{A,1})+p_1P^C_{z_A}(\uparrow|\lambda_1)N_{A,1})\big). \end{aligned} \tag{5.68}$$

And using (5.45),

$$\begin{aligned} &P^{QM}_{z_B,z_A}(\uparrow,\uparrow) = \\ &\quad P^C_{z_B}(\uparrow|\lambda_1)(\frac{4\frac{N^2_{B,1}}{N^2}-4\frac{N^3_{B,1}}{N^3}-\frac{N_{B,1}}{N}}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1)+2(\frac{N_{B,1}}{N}-\frac{N^2_{B,1}}{N^2})P^C_{z_A}(\uparrow|\lambda_2)) \\ &+P^C_{z_B}(\uparrow|\lambda_2)(2(\frac{N_{B,1}}{N}-\frac{N^2_{B,1}}{N^2})P^C_{z_A}(\uparrow|\lambda_2)+\frac{8\frac{N^2_{B,1}}{N^2}-4\frac{N^3_{B,1}}{N^3}-5\frac{N_{B,1}}{N}+1}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1)). \end{aligned} \tag{5.69}$$

We take the large N limit obtaining:

$$\begin{aligned} P^{QM}_{z_B,z_A}(\uparrow,\uparrow) &= P^C_{z_B}(\uparrow|\lambda_1)(\frac{4q^2_{B,1}-4q^3_{B,1}-q_{B,1}}{1-2q_{B,1}}P^C_{z_A}(\uparrow|\lambda_1)+2(q_{B,1}-q^2_{B,1})P^C_{z_A}(\uparrow|\lambda_2)) \\ &+ P^C_{z_B}(\uparrow|\lambda_2)(2(q_{B,1}-q^2_{B,1})P^C_{z_A}(\uparrow|\lambda_2)+\frac{8q^2_{B,1}-4q^3_{B,1}-5q_{B,1}+1}{1-2q_{B,1}}P^C_{z_A}(\uparrow|\lambda_1)). \end{aligned} \tag{5.70}$$

If we substitute the results derived from imposing the marginal probabilities, the expression becomes rather lengthy but simplifies to (see the third appendix for the details of the algebra), we write $q_{B1} = q$,

$$\begin{aligned} P^{QM}_{z_B,z_A}(\uparrow,\uparrow) &= \frac{(\frac{1}{2}-q)(P^C_{z_B}(\uparrow|\lambda_1)-(|\alpha|^2+|\gamma|^2))(P^C_{z_A}(\uparrow|\lambda_1)-(|\alpha|^2+|\beta|^2))}{q-1} \\ &+ (|\alpha|^2+|\beta|^2)(|\alpha|^2+|\gamma|^2). \end{aligned} \tag{5.71}$$

This should be equal to the probability predicted in QM, so

$$\begin{aligned} |\alpha|^2 &= \frac{(\frac{1}{2}-q)(P^C_{z_B}(\uparrow|\lambda_1)-(|\alpha|^2+|\gamma|^2))(P^C_{z_A}(\uparrow|\lambda_1)-(|\alpha|^2+|\beta|^2))}{q-1} \\ &+ (|\alpha|^2+|\beta|^2)(|\alpha|^2+|\gamma|^2). \end{aligned} \tag{5.72}$$

As a check, the expression resulting from the calculation if it had been done from Alice's frame can be calculated to be similar. Further, the equations following from the other joint probabilities can be shown to be the same equation, as can be seen when substituting $P^C_{z_B}(\uparrow|\lambda_1) = 1 - P^C_{z_B}(\downarrow|\lambda_1)$ and $|\alpha|^2 + |\gamma|^2 = 1 - |\delta|^2 - |\beta|^2$.

$$\begin{aligned}
|\alpha|^2 &= \frac{(\frac{1}{2} - q)(1 - P^C_{z_B}(\downarrow|\lambda_1) - (1 - |\delta|^2 - |\beta|^2))(P^C_{z_A}(\uparrow|\lambda_1) - (|\alpha|^2 + |\beta|^2))}{q-1} \\
&+ (|\alpha|^2 + |\beta|^2)(1 - |\delta|^2 - |\beta|^2), && (5.73) \\
|\alpha|^2 &= \frac{-(\frac{1}{2} - q)(P^C_{z_B}(\downarrow|\lambda_1) - (|\delta|^2 + |\beta|^2))(P^C_{z_A}(\uparrow|\lambda_1) - (|\alpha|^2 + |\beta|^2))}{q-1} \\
&+ |\alpha|^2 + |\beta|^2 - (|\alpha|^2 + |\beta|^2)(|\delta|^2 + |\beta|^2), && (5.74) \\
|\beta|^2 &= \frac{(\frac{1}{2} - q)(P^C_{z_B}(\downarrow|\lambda_1) - (|\delta|^2 + |\beta|^2))(P^C_{z_A}(\uparrow|\lambda_1) - (|\alpha|^2 + |\beta|^2))}{q-1} \\
&+ (|\alpha|^2 + |\beta|^2)(|\delta|^2 + |\beta|^2), && (5.75)
\end{aligned}$$

which is exactly the expression we would derive for $P^{QM}_{z_B,z_A}(\uparrow,\downarrow)$.

5.4.4 Checking particular examples

Let us investigate if this expression (5.72) could reproduce the probabilities of a very simple state, for example the triplet state

$$|1,1\rangle = |z\uparrow\rangle\,|z\uparrow\rangle\,, \tag{5.76}$$

so that in the above expressions for the coefficients, $|a| = 1$ and we obtain

$$\begin{aligned}
|\alpha|^2 &= \cos^2\frac{\theta_1}{2}\cos^2\frac{\theta_2}{2}, \\
|\beta|^2 &= \cos^2\frac{\theta_1}{2}\sin^2\frac{\theta_2}{2}, \\
|\gamma|^2 &= \sin^2\frac{\theta_1}{2}\cos^2\frac{\theta_2}{2}, \\
|\delta|^2 &= \sin^2\frac{\theta_1}{2}\sin^2\frac{\theta_2}{2}. && (5.77) \\
& && (5.78)
\end{aligned}$$

We deduce that (5.72) reduces to

$$0 = \frac{(\frac{1}{2} - q)(P^C_{z_A}(\uparrow|\lambda_1) - cos^2\frac{\theta_1}{2})(P^C_{z_B}(\uparrow|\lambda_1) - cos^2\frac{\theta_2}{2})}{q-1}. \tag{5.79}$$

We are now forced to ask ourselves what the role of q is in this whole approach. Can it depend on the angles? For more complicated states we will see that, if the whole approach should make sense, it must have angle dependence. Yet for this example it seems it does not matter what q is, it is enough that

$$P^C_{z_A}(\uparrow|\lambda_1) = \cos^2\frac{\theta_1}{2} \tag{5.80}$$

Using the expressions (5.62) and (5.65) and requiring the other conditional probabilities to take admissible values between 0 and 1, it becomes clear that it is convenient, if q should remain to be chosen freely, that

$$P^C_{z_B}(\uparrow|\lambda_1) = \cos^2\frac{\theta_2}{2}, \tag{5.81}$$

and so

$$P^C_{z_A}(\uparrow|\lambda_2) = \cos^2\frac{\theta_1}{2}\frac{q-1}{q-1} = \cos^2\frac{\theta_1}{2} \qquad P^C_{z_B}(\uparrow|\lambda_2) = \cos^2\frac{\theta_2}{2}. \tag{5.82}$$

We conclude that the conditional probabilities given λ_1 are exactly the same as those for given λ_2. Apparently, for this state, no distinction is made between the two values the hidden variable could take, both for Alice and Bob. One could consider this somewhat dissatisfying. This was supposed to be the variety which enables us to say more about the situation than QM would. So let us instead try $q = 1/2$. (5.62) then becomes

$$P^C_{z_A}(\uparrow|\lambda_2) = \frac{-\frac{\cos^2\theta_1/2}{4}}{-1/4} = \cos^2\frac{\theta_1}{2}. \tag{5.83}$$

$P^C_{z_A}(\uparrow|\lambda_1)$ can then be chosen different. Taking $q = 1/2$ we see that (5.65) still requires

$$P^C_{z_B}(\uparrow|\lambda_1) = \cos^2\frac{\theta_2}{2}. \tag{5.84}$$

So it seems for this state that we can either choose to fix the conditional probabilities, obtaining freedom to choose q, or we choose to fix the source probability q, obtaining some freedom to choose the conditional probabilities of one of the parties.

Let us try the more advanced example of the singlet. We then have the QM state:

$$|0,0\rangle = \frac{1}{\sqrt{2}}|z\uparrow\rangle_A \otimes |z\downarrow\rangle_B - \frac{1}{\sqrt{2}}|z\downarrow\rangle_A \otimes |z\uparrow\rangle_B\,. \tag{5.85}$$

So in (5.51) we simply take $|b| = |c| = 1/\sqrt{2}$ and $|a| = |d| = 0$

$$\begin{aligned}
|\alpha|^2 &= \frac{1}{2}\cos^2\frac{\theta_1}{2}\sin^2\frac{\theta_2}{2} + \frac{1}{2}\cos^2\frac{\theta_2}{2}\sin^2\frac{\theta_1}{2} - \cos(\phi_2-\phi_1)\cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2}\sin\frac{\theta_2}{2}\sin\frac{\theta_1}{2}, \\
|\beta|^2 &= \frac{1}{2}\cos^2\frac{\theta_1}{2}\cos^2\frac{\theta_2}{2} + \frac{1}{2}\sin^2\frac{\theta_2}{2}\sin^2\frac{\theta_1}{2} + \cos(\phi_2-\phi_1)\cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2}\sin\frac{\theta_2}{2}\sin\frac{\theta_1}{2}, \\
|\gamma|^2 &= \frac{1}{2}\sin^2\frac{\theta_1}{2}\sin^2\frac{\theta_2}{2} + \frac{1}{2}\cos^2\frac{\theta_2}{2}\cos^2\frac{\theta_1}{2} + \cos(\phi_2-\phi_1)\cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2}\sin\frac{\theta_2}{2}\sin\frac{\theta_1}{2}.
\end{aligned} \tag{5.86}$$

So we see immediately that the marginal probabilities are always 1/2. We see that, after some algebra, (5.72) becomes

$$\begin{aligned}
&-\frac{1}{4}\cos\theta_1\cos\theta_2 - \frac{1}{4}\cos(\phi_2-\phi_1)\sin\theta_2\sin\theta_1 + \frac{1}{4} \\
&\qquad = \frac{(\frac{1}{2}-q)(P^C_{z_B}(\uparrow|\lambda_1)-\frac{1}{2})(P^C_{z_A}(\uparrow|\lambda_1)-\frac{1}{2})}{q-1} + \frac{1}{4},
\end{aligned} \tag{5.87}$$

$$\cos\theta_1\cos\theta_2+\cos(\phi_2-\phi_1)\sin\theta_2\sin\theta_1 = \frac{2(2q-1)(P^C_{z_B}(\uparrow|\lambda_1)-\frac{1}{2})(P^C_{z_A}(\uparrow|\lambda_1)-\frac{1}{2})}{q-1}. \tag{5.88}$$

From this expression it becomes clear that q cannot just be a constant if the conditional probability of Alice should only depend on her angles θ_1 and ϕ_1 and Bob's only on θ_1 and ϕ_1. If that would be the case, one could derive an expression for $P^C_{z_A}(\uparrow|\lambda_1)(\theta_1,\phi_1)$ setting θ_2 and ϕ_2 zero and an expression for $P^C_{z_B}(\uparrow|\lambda_1)(\theta_1,\phi_1)$ setting θ_1 and ϕ_1 zero. Then substituting these expressions back in (5.88) gives a dependence on the ϕ on the left hand side, but not on the right hand side. It seems therefore that we have to either assume the conditional probabilities to depend on all angles, or we cannot simply take q to be a constant. We will go with the second option here. This seems more acceptable, as we supposed the lambdas do not look the same for observers in different frames (depending on the frame), so does the frequency of their appearance. So let us see, still using the previously explained strategy, what expression we obtain for $q(\theta_1,\theta_2,\phi_1,\phi_2)$. Setting θ_2 and ϕ_2 zero, using short hand expressions $P_{A,1}(\theta_1,\phi_1)$ for Alice's conditional probability to measure spin up given a λ_1 (and likewise for Bob),

$$\cos\theta_1\frac{q(\theta_1,\phi_1,0,0)-1}{2(2q(\theta_1,\phi_1,0,0)-1)} = (P_{B,1}(0,0)-\frac{1}{2})(P_{A,1}(\theta_1,\phi_1)-\frac{1}{2}). \tag{5.89}$$

So we deduce

$$P_{A,1}(\theta_1,\phi_1) = \cos\theta_1 \frac{q(\theta_1,\phi_1,0,0)-1}{2(P_{B,1}(0,0)-\frac{1}{2})(2q(\theta_1,\phi_1,0,0)-1)} + \frac{1}{2}, \tag{5.90}$$

$$P_{B,1}(\theta_2,\phi_2) = \cos\theta_2 \frac{q(\theta_2,\phi_2,0,0)-1}{2(P_{A,1}(0,0)-\frac{1}{2})(2q(\theta_2,\phi_2,0,0)-1)} + \frac{1}{2}, \tag{5.91}$$

and

$$\frac{q(0,0,0,0)-1}{2(2q(0,0,0,0)-1)} = (P_{B,1}(0,0)-\frac{1}{2})(P_{A,1}(0,0)-\frac{1}{2}). \tag{5.92}$$

Substituting these back in (5.88) and rearranging

$$\cos\theta_1\cos\theta_2 + \cos(\phi_2-\phi_1)\sin\theta_2\sin\theta_1 = \\ \frac{(2q(\theta_1,\theta_2,\phi_1,\phi_2)-1)\cos\theta_1\cos\theta_2}{q(\theta_1,\theta_2,\phi_1,\phi_2)-1}\frac{2q(0,0,0,0)-1}{q(0,0,0,0)-1}\frac{q(\theta_2,\phi_2,0,0)-1}{2q(\theta_2,\phi_2,0,0)-1}\frac{q(\theta_1,\phi_1,0,0)-1}{2q(\theta_1,\phi_1,0,0)-1}, \tag{5.93}$$

which would be quite a challenge to solve. If we make the simplifying assumption that $P_{A,1}(\theta_1,\phi_1) = P_{A,1}(\theta_1) = \cos^2\theta_1/2$, as it was for $|1,0\rangle$, we obtain that

$$\cos^2\frac{\theta_1}{2} - \frac{1}{2} = \cos\theta_1 \frac{q(\theta_1,\phi_1,0,0)-1}{2(P_{B,1}(0,0)-\frac{1}{2})(2q(\theta_1,\phi_1,0,0)-1)}, \tag{5.94}$$

$$1 = \frac{q(\theta_1,\phi_1,0,0)-1}{(P_{B,1}(0,0)-\frac{1}{2})(2q(\theta_1,\phi_1,0,0)-1)}. \tag{5.95}$$

This time we take $P_{B,1}(\theta_1) = \sin^2\theta_2/2$, so that

$$\sin^2\frac{\theta_1}{2} - \frac{1}{2} = -\frac{1}{2}\cos\theta_1 = \cos\theta_1 \frac{q(\theta_1,\phi_1,0,0)-1}{2(P_{A,1}(0,0)-\frac{1}{2})(2q(\theta_1,\phi_1,0,0)-1)}, \tag{5.96}$$

$$-1 = \frac{q(\theta_2,\phi_2,0,0)-1}{(P_{A,1}(0,0)-\frac{1}{2})(2q(\theta_2,\phi_2,0,0)-1)}. \tag{5.97}$$

(5.88) Would then become

$$\cos\theta_1\cos\theta_2 + cos(\phi_2-\phi_1)\sin\theta_2\sin\theta_1 = -\frac{(q(\theta_1,\theta_2,\phi_1,\phi_2)-\frac{1}{2})\cos\theta_1\cos\theta_2}{q(\theta_1,\theta_2,\phi_1,\phi_2)-1}. \tag{5.98}$$

So we would obtain

$$q(\theta_1,\theta_2,\phi_1,\phi_2) = \frac{3/2+\cos(\phi_2-\phi_1)\tan\theta_2\tan\theta_1}{2+\cos(\phi_2-\phi_1)\tan\theta_2\tan\theta_1}. \tag{5.99}$$

This expression for $q(\theta_1, \theta_2, \phi_1, \phi_2)$ can be checked to satisfy (5.95) and (5.97) for these choices of $P_{A,1}(\theta_1)$ and $P_{B,1}(\theta_2)$, as it should. Yet this function $q(\theta_1, \theta_2, \phi_1, \phi_2)$ can be hardly thought of to be a probability, as it becomes negative or greater than 1 for some particular combination of angles. Yet for $\theta_1(\theta_2)$ zero, which is what Alice(Bob) would do describing the situation from their point of view, $q = 3/4$ and (5.88) produces the right probability, so all seems right. Yet if we, without letting this fact bother us, simply take $P_{A,1}(\theta_1) = cos^2\theta_1/2$ and $P_{B,1}(\theta_2) = sin^2\theta_2$ and set θ_1 zero, and so $q = 3/4$, then it is not hard to derive from (5.62) and (5.65) that it should count

$$P^C_{z_A}(\uparrow|\lambda_2) = \frac{1}{3}, \qquad P^C_{z_A}(\downarrow|\lambda_2) = \frac{2}{3}. \tag{5.100}$$

The other probability causes more trouble

$$P^C_{z_B}(\uparrow|\lambda_2) = -3sin^2\theta_2/2 + 2, \qquad P^C_{z_B}(\downarrow|\lambda_2) = -3cos^2\theta_2/2 + 2, \tag{5.101}$$

which are only an acceptable quantities for a probability if $2\arccos[\sqrt{2/3}] \approx 1.23 < \theta_2 < 1.91 \approx 2\arcsin\sqrt{2/3}$ radians. We should therefore be modest about the range of validity of this model.

The triplet state

$$|1,0\rangle = \frac{1}{\sqrt{2}}\,|z\uparrow\rangle_A \otimes |z\downarrow\rangle_B + \frac{1}{\sqrt{2}}\,|z\downarrow\rangle_A \otimes |z\uparrow\rangle_B\,, \tag{5.102}$$

works quite similar to the singlet state. The terms in the joint probabilities only differ by some minus signs

$$\begin{aligned}
|\alpha|^2 &= \frac{1}{2}\cos^2\frac{\theta_1}{2}\sin^2\frac{\theta_2}{2} + \frac{1}{2}\cos^2\frac{\theta_2}{2}\sin^2\frac{\theta_1}{2} + \cos(\phi_2-\phi_1)\cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2}\sin\frac{\theta_2}{2}\sin\frac{\theta_1}{2},\\
|\beta|^2 &= \frac{1}{2}\cos^2\frac{\theta_1}{2}\cos^2\frac{\theta_2}{2} + \frac{1}{2}\sin^2\frac{\theta_2}{2}\sin^2\frac{\theta_1}{2} - \cos(\phi_2-\phi_1)\cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2}\sin\frac{\theta_2}{2}\sin\frac{\theta_1}{2},\\
|\gamma|^2 &= \frac{1}{2}\sin^2\frac{\theta_1}{2}\sin^2\frac{\theta_2}{2} + \frac{1}{2}\cos^2\frac{\theta_2}{2}\cos^2\frac{\theta_1}{2} - \cos(\phi_2-\phi_1)\cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2}\sin\frac{\theta_2}{2}\sin\frac{\theta_1}{2}.
\end{aligned} \tag{5.103}$$

So we see immediately that the marginal probabilities are always 1/2. We see that, after some algebra, (5.72) becomes

$$\cos\theta_1\cos\theta_2 - \cos(\phi_2-\phi_1)\sin\theta_2\sin\theta_1 = \frac{2(2q-1)(P^C_{z_B}(\uparrow|\lambda_1)-\frac{1}{2})(P^C_{z_A}(\uparrow|\lambda_1)-\frac{1}{2})}{q-1}. \tag{5.104}$$

We can basically repeat the previous strategy we used with the singlet state, obtaining for the same assumptions for $P_{A,1}(\theta_1, \phi_1)$ and $P_{B,1}(\theta_2)$ that

$$q(\theta_1, \theta_2, \phi_1, \phi_2) = \frac{3/2 - \cos(\phi_2 - \phi_1)\tan\theta_2 \tan\theta_1}{2 - \cos(\phi_2 - \phi_1)\tan\theta_2 \tan\theta_1}. \tag{5.105}$$

It seems we run in similar trouble as with the singlet: q does not seem to be interpretable as a probability for all angles. Yet again, there seems to be no trouble for $1.23 < \theta_2 < 1.91$ and θ_1 zero. Then we simply have $q = 3/4$ and the joint probability takes a value both Alice and Bob would calculate in their own frame.

These results are of course far from satisfactory and our expectation is that a better solution should be possible. Perhaps the improvement only appears when more than two values for the hidden variable is considered, which gives the model far more degrees of freedom. This is a subject for future research.

5.5 Frame Transformations

In the previous section the condition for Alice and Bob to agree on all joint probabilities if the information about λ fixed the values of the probabilities p_n. These results could have been derived more easily if the expression for the joint probabilities was interpreted as in-product between two vectors

$$\begin{pmatrix} P_{z_B}(\uparrow|\lambda_1) \\ P_{z_B}(\uparrow|\lambda_2) \end{pmatrix} = \vec{v}_{z_B,\uparrow} \qquad \begin{pmatrix} P_{z_A}(\uparrow|\lambda_1)^B \\ P_{z_A}(\uparrow|\lambda_2^B) \end{pmatrix} = \vec{v}^{\,B}_{z_A,\uparrow}$$

where the product is taken using a metric, defined in Bob's frame as

$$\eta_{z_B} = \begin{pmatrix} q_{B,1} & \\ & 1 - q_{B,1} \end{pmatrix}.$$

In Alice's frame the metric is

$$\eta_{z_A} = \begin{pmatrix} q_{A,1} & \\ & 1 - q_{A,1} \end{pmatrix}.$$

The transformation matrix for the vectors (between Alice and Bob's frame) was written in the matrix form introduced in section (5.3.2). As usual, we let $\vec{v}^T$ denote the transpose of $\vec{v}$. The condition for the joint probability for both Alice and Bob to measure spin up could then be written, if $\vec{v}^B_{z_A,\uparrow} = M_{AB}\vec{v}_{z_A,\uparrow}$ and $\vec{v}^A_{z_B,\uparrow} = M_{BA}\vec{v}_{z_B,\uparrow}$,

$$\begin{aligned} (\vec{v}^B_{z_A,\uparrow})^T\eta_{z_B}\vec{v}_{z_B,\uparrow} &= \vec{v}^T_{z_A,\uparrow}\eta_{z_A}\vec{v}^A_{z_B,\uparrow}, && (5.106) \\ \vec{v}^T_{z_A,\uparrow}M^T_{AB}\eta_{z_B}\vec{v}_{z_B,\uparrow} &= \vec{v}^T_{z_A,\uparrow}\eta_{z_A}M_{BA}\vec{v}_{z_B,\uparrow}, && (5.107) \\ \vec{v}^T_{z_A,\uparrow}(M^T_{AB}\eta_{z_B} - \eta_{z_A}M_{BA})\vec{v}_{z_B,\uparrow} &= 0 && (5.108) \end{aligned}$$

This should count for any $\vec{v}$, so the condition reduces to

$$\begin{aligned} M^T_A\eta_{z_B} - \eta_{z_A}M_B &= 0, && (5.109) \\ \eta_{z_B} &= (M^T_{AB})^{-1}\eta_{z_A}M_{BA}. && (5.110) \end{aligned}$$

Using the requirements (5.31)

$$\eta_{z_B} = M^T_{BA}\eta_{z_A}M_{BA}. \tag{5.111}$$

In Special Relativity a similar transformation is rule is defined for the metric. It defines actually the matrices M_{BA} as Lorentz transformations. In Special Relativity the Lorentz transformation describes how space and time intervals differ between two different observers. In this PHV model, it describes how probabilities differ to observers in different frames. The analogy between the PHV model and Special Relativity at the beginning of this chapter seems therefore in place.

Note that (5.111) is automatically satisfied with (5.45). It is therefore quite surprising how the requirement of agreement between Alice and Bob about the joint probabilities automatically leads to the identification of the transformation between frames with a Lorentz transformation. If we take the reverse relation, the transformation law of the metric may be the reason why agreement is reached about joint probabilities, and therefore why QM produces the probabilities it produces (if this PHV model could be taken serious). It is not exactly clear how the identified metric in this PHV should be interpreted.

5.6 Generalization to multiple and continuous λs

5.6.1 Three λs

We will now see how the previous model for two λs can be generalized to any number of λs, or even an infinite number of them. In order for a difference to occur, we needs the transformation $\lambda_{B,1} \rightarrow \lambda^A_{B,1}$ to mix up the two subsets for Bob. Let us not be too ambitious first and see how the relations change if instead of two there will be three λs, λ_1,λ_2 and λ_3. Each of them defines a subset in the measurement results of the EPR Bohm set up which could be different for Bob and Alice. Similar to the two λ case we can define the transformation rules in terms of probabilities:

$$\Lambda(\lambda_1) = \begin{cases} \lambda_2 & \text{with probability } P_{1,2} \\ \lambda_3 & \text{with probability } P_{1,3} \\ \lambda_1 & \text{with probability } 1 - P_{1,2} - q_{1,3} \end{cases}$$

$$\Lambda(\lambda_2) = \begin{cases} \lambda_1 & \text{with probability } P_{2,1} \\ \lambda_3 & \text{with probability } P_{2,3} \\ \lambda_2 & \text{with probability } 1 - P_{2,1} - q_{2,3} \end{cases}$$

$$\Lambda(\lambda_3) = \begin{cases} \lambda_1 & \text{with probability } P_{3,1} \\ \lambda_2 & \text{with probability } P_{3,2} \\ \lambda_3 & \text{with probability } 1 - P_{3,1} - q_{3,2} \end{cases}$$

Again it would be more convenient to define fractions as our variables. Similar to the previous sections, Alice could define $p_{1,2}$ as the fraction of $N_{1,B}$ representing the number of λ_1s she turns into λ_2s. Similarly, it tells Alice to, not looking at Bob's spin values, change an arbitrary fraction $p_{1,3}$ of the λ_1 cases to λ_3. Similar to before the transformed number of spin ups in the subset defined by λ_1 then becomes

$$n^A_{\uparrow,B,1} = (1 - p_{1,2} - p_{1,3})n_{\uparrow,B,1} + p_{2,1}n_{\uparrow,B,2} + p_{3,1}n_{\uparrow,B,3}. \tag{5.112}$$

It is convenient to write down the transformation of a vector of the populations $n_{\uparrow,B,1}$,$n_{\uparrow,B,2}$ and $n_{\uparrow,B,3}$ (and similar for the number of spin downs) in a matrix form as

$$\begin{pmatrix} n^A_{\uparrow,B,1} \\ n^A_{\uparrow,B,2} \\ n^A_{\uparrow,B,3} \end{pmatrix} = \begin{pmatrix} 1 - p_{1,2} - p_{1,3} & p_{2,1} & p_{3,1} \\ p_{1,2} & 1 - p_{2,1} - p_{2,3} & p_{3,2} \\ p_{1,3} & p_{2,3} & 1 - p_{3,1} - p_{3,2} \end{pmatrix} \begin{pmatrix} n_{\uparrow,B,1} \\ n_{\uparrow,B,2} \\ n_{\uparrow,B,2} \end{pmatrix},$$

and similarly for a vector with $n_{\downarrow,1,B}$,$n_{\downarrow,2,B}$ and $n_{\downarrow,3,B}$. The underlying assumption is again that the number of $n_\uparrow$ is reduced by the same factor as $n_\downarrow$ in the subset characterized by $\lambda_{1,B}$ so that Alice's adjustments are completely arbitrary. The transformation can be inverted to find out how Bob would transform Alice's probabilities. This gives a huge expression in terms of the ps defined above which does not add to much meaning, so we will not present it here. The next step would be to define how a vector with the conditional probabilities $P^C_{z_B}(\uparrow|\lambda_n)$ transform. As

$P^C_{z_B}(\uparrow|\lambda_1)^A = n^A_\uparrow/N^A_{B,1}$

$$P^C_{z_B}(\uparrow|\lambda_1)^A = (1-p_{1,2}-p_{1,3})P^C_{z_B}(\uparrow|\lambda_{B,1})\frac{N_{B,1}}{N^A_{B,1}}+p_{2,1}P^C_{z_B}(\uparrow|\lambda_{B,2})\frac{N_{B,2}}{N^A_{B,1}}+p_{3,1}P^C_{z_B}(\uparrow|\lambda_{B,3})\frac{N_{B,3}}{N^A_{B,1}}, \tag{5.113}$$

where $N^A_{B,1}$ is the size of Bob's subset defined by λ_1 as Alice perceives it (so, the size of the transformed subset). It is given by

$$\begin{aligned} N^A_{B,1} &= n^A_{\uparrow,B,1} + n^A_{\downarrow,B,1} \\ &= (1-p_{1,2}-p_{1,3})n_{\uparrow,B,1} + p_{2,1}n_{\uparrow,B,2} + p_{3,1}n_{\uparrow,B,3} \\ &+ (1-p_{1,2}-p_{1,3})n_{\downarrow,B,1} + p_{2,1}n_{\downarrow,B,2} + p_{3,1}n_{\downarrow,B,3} \\ &= (1-p_{1,2}-p_{1,3})N_{B,1} + p_{2,1}N_{B,2} + p_{3,1}N_{B,3}. \end{aligned} \tag{5.114}$$

We also have the requirement

$$N = N_{B,1} + N_{B,2} + N_{B,3} = N^A_{B,1} + N^A_{B,2} + N^A_{B,3} \tag{5.115}$$

Which should also be satisfied for Alice's $N_{A,1}, N^B_{A,1}, ...$ Writing the transformation in a matrix form again.

$$\begin{pmatrix} P_{z_B}(\uparrow|\lambda^A_1) \\ P_{z_B}(\uparrow|\lambda^A_2) \\ P_{z_B}(\uparrow|\lambda^A_3) \end{pmatrix} =$$

$$\begin{pmatrix} (1-p_{1,2}-p_{1,3})\frac{N_{B,1}}{N^A_{B,1}} & p_{2,1}\frac{N_{B,2}}{N^A_{B,1}} & p_{3,1}\frac{N_{B,3}}{N^A_{B,1}} \\ p_{1,2}\frac{N_{B,1}}{N^A_{B,2}} & (1-p_{2,1}-p_{2,3})\frac{N_{B,2}}{N^A_{B,2}} & p_{3,2}\frac{N_{B,3}}{N^A_{B,2}} \\ p_{1,3}\frac{N_{B,1}}{N^A_{B,3}} & p_{2,3}\frac{N_{B,2}}{N^A_{B,3}} & (1-p_{3,1}-p_{3,2})\frac{N_{B,3}}{N^A_{B,3}} \end{pmatrix} \begin{pmatrix} P_{z_B}(\uparrow|\lambda_1) \\ P_{z_B}(\uparrow|\lambda_2) \\ P_{z_B}(\uparrow|\lambda_3) \end{pmatrix}.$$

The inverse of this transformation should describe how Bob sees Alice's probabilities transform. Again because this calculation would take up a lot of space we refer to appendix 2 for the details. It turns out indeed the transformation is the inverse if the given conditions are met:

$$N^A_{B,1} = N_{A,1}, \qquad N^A_{B,2} = N_{A,2}, \qquad N^A_{B,3} = N_{A,3}. \tag{5.116}$$

Again we can define Alice and Bob's metric

$$\eta_{z_A} = \begin{pmatrix} \frac{N_{A,1}}{N} & & \\ & \frac{N_{A,2}}{N} & \\ & & \frac{N_{A,3}}{N} \end{pmatrix} \qquad \eta_{z_B} = \begin{pmatrix} \frac{N_{B,1}}{N} & & \\ & \frac{N_{B,2}}{N} & \\ & & \frac{N_{B,3}}{N} \end{pmatrix},$$

which we see are in accordance with the expression for the joint probability in QM for both to measure up, according to Alice

$$
\begin{aligned}
&P^{J,QM}(\uparrow,\uparrow) = \\
&P^C_{z_A}(\uparrow|\lambda_1)P^C_{z_B}(\uparrow|\lambda_1)^A\frac{N_{A,1}}{N}+P^C_{z_A}(\uparrow|\lambda_2)P^C_{z_B}(\uparrow|\lambda_2)^A\frac{N_{A,2}}{N}+P^C_{z_A}(\uparrow|\lambda_3)P^C_{z_B}(\uparrow|\lambda_3')\frac{N_{A,3}}{N}.
\end{aligned}
\tag{5.117}
$$

Setting this equal again to what Bob would calculate is the QM joint probability (and similarly for the other joint probabilities) effectively comes down to satisfying

$$(\vec{v}^B_{z_A})^T\eta_{z_B}\vec{v}_{z_B} = \vec{v}^T_{z_A}\eta_{z_A}\vec{v}^A_{z_B}, \tag{5.118}$$

$$\vec{v}^T_{z_A}(M^T_{AB}\eta_{z_B} - \eta_{z_A}M_{BA})\vec{v}_{z_B} = 0, \tag{5.119}$$

$$
\begin{pmatrix} \frac{N_{B,1}}{N} & & \\ & \frac{N_{B,2}}{N} & \\ & & \frac{N_{B,3}}{N} \end{pmatrix} = \begin{pmatrix} (1-p_{1,2}-p_{1,3})\frac{N_{B,1}}{N^A_{B,1}} & p_{1,2}\frac{N_{B,1}}{N^A_{B,2}} & p_{1,3}\frac{N_{B,1}}{N^A_{B,3}} \\ p_{2,1}\frac{N_{B,2}}{N^A_{B,1}} & (1-p_{2,1}-p_{2,3})\frac{N_{B,2}}{N^A_{B,2}} & p_{2,3}\frac{N_{B,2}}{N^A_{B,3}} \\ p_{3,1}\frac{N_{B,3}}{N^A_{B,1}} & p_{3,2}\frac{N_{B,3}}{N^A_{B,2}} & (1-p_{3,1}-p_{3,2})\frac{N_{B,3}}{N^A_{B,3}} \end{pmatrix},
$$

$$
\cdot\begin{pmatrix} \frac{N_{A,1}}{N} & & \\ & \frac{N_{A,2}}{N} & \\ & & \frac{N_{A,3}}{N} \end{pmatrix} \begin{pmatrix} (1-p_{1,2}-p_{1,3})\frac{N_{B,1}}{N^A_{B,1}} & p_{2,1}\frac{N_{B,2}}{N^A_{B,1}} & p_{3,1}\frac{N_{B,3}}{N^A_{B,1}} \\ p_{1,2}\frac{N_{B,1}}{N^A_{B,2}} & (1-p_{2,1}-p_{2,3})\frac{N_{B,2}}{N^A_{B,2}} & p_{3,2}\frac{N_{B,3}}{N^A_{B,2}} \\ p_{1,3}\frac{N_{B,1}}{N^A_{B,3}} & p_{2,3}\frac{N_{B,2}}{N^A_{B,3}} & (1-p_{3,1}-p_{3,2})\frac{N_{B,3}}{N^A_{B,3}} \end{pmatrix}.
$$

This results in the following equations for the different matrix elements η_{nm}

$$\eta_{11} = \frac{1}{N_{B,1}} = \frac{(1-p_{1,2}-p_{1,3})^2}{N^A_{B,1}} + \frac{p^2_{1,2}}{N^A_{B,2}} + \frac{p^2_{1,3}}{N^A_{B,3}}, \tag{5.120}$$

$$\eta_{22} = \frac{1}{N_{B,2}} = \frac{p^2_{2,1}}{N^A_{B,1}} + \frac{(1-p_{2,1}-p_{2,3})^2}{N^A_{B,2}} + \frac{p^2_{2,3}}{N^A_{B,3}}, \tag{5.121}$$

$$\eta_{33} = \frac{1}{N_{B,3}} = \frac{p^2_{3,1}}{N^A_{B,1}} + \frac{p^2_{3,2}}{N^A_{B,2}} + \frac{(1-p_{3,1}-p_{3,2})^2}{N^A_{B,3}}, \tag{5.122}$$

$$\eta_{12} = 0 = \frac{(1-p_{1,2}-p_{1,3})p_{2,1}}{N^A_{B,1}} + \frac{(1-p_{2,1}-p_{2,3})p_{1,2}}{N^A_{B,2}} + \frac{p_{2,3}p_{1,3}}{N^A_{B,3}}, \tag{5.123}$$

$$\eta_{13} = 0 = \frac{(1-p_{1,2}-p_{1,3})p_{3,1}}{N^A_{B,1}} + \frac{p_{3,2}p_{1,2}}{N^A_{B,2}} + \frac{p_{1,3}(1-p_{3,1}-p_{3,2})}{N^A_{B,3}}, \tag{5.124}$$

$$\eta_{23} = 0 = \frac{p_{3,1}p_{2,1}}{N^A_{B,1}} + \frac{(1-p_{2,1}-p_{2,3})p_{3,2}}{N^A_{B,2}} + \frac{p_{2,3}(1-p_{3,1}-p_{3,2})}{N^A_{B,3}}. \tag{5.125}$$

It can be seen immediately that the equations for the η_{nm} matrix elements are the same as the one for the η_{mn} matrix element. We therefore conclude there are only 6 independent equations. That is just enough to fix 6 variables $p_{n,m}$ (for $n \neq m$) in terms of N, $N_{B,1}$ and $N_{B,2}$, (for $N_{B,3} = N - N_{B,1} - N_{B,2}$). It turns out to be very tedious to solve these equations though and it is advised to use a computational device for it. Our attempts so far did not give any clear results.

5.6.2 Multiple and continuous λs

It may be clear that this story could be generalized easily to four, five or a greater number of different λs. It may also be clear that it quickly becomes a lot of work solving the equations for the different probabilities: for n λs, the transformation matrices become nxn, which means there is $n(n+1)/2$ equations to solve for $n > 2$. On the other hand, we observer that for n λ's, the number of probabilities p_n defining the transformation of the λs is $n(n-1)$. It seems there will be, as n increases, more and more freedom to fix the values p_n. We can now also better understand why only the trivial case was a solution for $n = 2$.)

We will now investigate if it is also possible to go to a continuous λ defined in a certain domain Λ. Obviously the number of possible different λs is then always bigger than N, although N is always chosen very large. The transformation of the λs would have to be defined differently. One could define a function $\rho(\lambda_{z_A}, \lambda)$ which is a specific probability distribution for a given λ_{z_A}. Integrated over a certain interval it gives the probability (as a function of λ_{z_A}) for λ_{z_A} to transform into a λ which is inside this interval. Let us normalize the probability distribution:

$$\int_\Lambda \rho(\lambda_{z_A}, \lambda) d\lambda = 1. \tag{5.126}$$

If Alice then measures a value of a λ_{z_A}, she knows Bob measures a value $\lambda_{z_B} = \lambda'_{z_A}$ (assuming again that if they were in the same frame, $\lambda_{z_A} = \lambda_{z_B}$). Although there is no way for Alice to know how her value transformed with certainty, as the transformation is probabilistic, the expectation value $\langle \lambda'_{z_A} \rangle$ would be a good approximation for a large N.

$$\langle \lambda'_{z_A} \rangle = \int_\Lambda \lambda \rho(\lambda_{z_A}, \lambda) d\lambda. \tag{5.127}$$

So Alice can for every λ_{z_A} she has go to Bob's results and change the value λ_{z_B} Bob has for that trial into $\langle \lambda'_{z_A} \rangle$. Bob can do the same with Alice's values using an probability distribution $\rho^{-1}(\lambda_{z_B}, \lambda)$ which can be considered the inverse of Alice's transformation in the sense that

$$\int_\Lambda \rho(\lambda_{z_A}, \lambda)\rho^{-1}(\lambda, \lambda_{z_B})d\lambda = \delta(\lambda_{z_A} - \lambda_{z_B}). \tag{5.128}$$

Alice can then define a series of n intervals $\Delta\lambda_n$ such that

$$\sum_n \Delta\lambda_n \tag{5.129}$$

covers the domain Λ completely. She could then go over her own list and determine the probabilities to measure spin up $P^C_{z_A}(\uparrow|\lambda \in \Delta\lambda_n)$ for different n. Likewise she could determine Bob's conditional probabilities $P^C_{z_B}(\uparrow|\lambda' \in \Delta\lambda_n)$ on the basis of the transformed values of λ.

Alice could have changed the λs of Bob also using these domains. She could have defined the intervals $\Delta\lambda_n$ first and then calculate the probability for a given λ_{z_A} to be in one of those particular intervals. This she could do for many different λ_{z_A}s. Then she could make subsections of her own data on the basis of the intervals (every λ_{z_A} belongs to one interval uniquely). Then, working within one of those subsections, she could consider the individual cases on Bob's side and sort them in the different domains. Without any further information about the transformation law it is unclear which way would be preferable. Further investigation is therefore needed for this case of a continuous λ.

5.7 Generalization to multipartite states

So far we have only considered bipartite states, which involves two observers. If this PHV model is any good, it should also generalize to the cases with more than two observers. Let us suppose we have three spin 1/2 particles on which we only do spin measurements. For now we will consider a discrete PHV having only a λ_1 and λ_2. We introduce a third observer Cherique who is doing measurements on her part of a tripartite state. The transformations between Alice and Bob described in the previous chapter can be taken over automatically from the previous chapter, defining the probabilities $P_{n,AB}$. A different transformation applies between the frames of Alice and Cherique, defining the $P_{n,AC}$. Another 'new' parameter entering is a $N_{C,1}$ describing Cherique's different populations of the λ_1 and λ_2. The relation between Bob and Cherique is described similarly, with yet another transformation, defining a $P_{1,BC}$ and a $P_{2,CB}$. Of course Cherique, as an independent observer, also has her own metric η_C.

Knowing that the requirements imposed by the joint probabilities fix both probabilities for each transformation in terms of $N_{A/B/C,1}$, it seems no further problems should arise in this model. Something which may be slightly different from the case with two observers is the requirement (these can be shown to be really two equations, fixing $N_{B,1}$ and $N_{C,1}$ in terms of $N_{A,1}$):

$$N_{A,1} = N^A_{B,1} = N^A_{C,1}, \qquad N^C_{A,1} = (N^A_{B,1})^C = N_{C,1}, \qquad N^B_{A,1} = N_{B,1} = (N^A_{C,1})^B. \tag{5.130}$$

The equations basically established the symmetry of the transformations between Alice and Bob and Alice and Cherique. Of course a similar relation should count between Bob and Cherique, which is could then be interpreted as the two other transformations one after the other (note they do not necessarily commute). This suggest a group structure to the transformation.

What is new in this approach for the tripartite state is the requirement for all three parties to agree on the joint probability (say, for all three to measure up) as QM predicts it. The structure of the transformations implements them to agree in pairs of two to agree on joint probabilities concerning only that pair, but not yet (at least clearly) for the case of three at the same time. From Alice's frame this joint probability would be given by

$$P^{J,QM}(\uparrow,\uparrow,\uparrow) =$$
$$P^C_{z_A}(\uparrow|\lambda_1)P^C_{z_B}(\uparrow|\lambda_1)^A P^C_{z_C}(\uparrow|\lambda_1)^A\frac{N_{A,1}}{N}+P^C_{z_A}(\uparrow|\lambda_2)P^C_{z_B}(\uparrow|\lambda_2)^A P^C_{z_C}(\uparrow|\lambda_2)^A\frac{N-N_{A,1}}{N}. \tag{5.131}$$

Bob would say it is

$$P^C_{z_A}(\uparrow|\lambda_1)^B P^C_{z_B}(\uparrow|\lambda_1)P^C_{z_C}(\uparrow|\lambda_1)^B\frac{N_{B,1}}{N}+P^C_{z_A}(\uparrow|\lambda_2)^B P^C_{z_B}(\uparrow|\lambda_2)P^C_{z_C}(\uparrow|\lambda_2)^B\frac{N-N_{B,1}}{N}. \tag{5.132}$$

Finally, Cherique would claim:

$$P^C_{z_A}(\uparrow|\lambda_1)^C P^C_{z_B}(\uparrow|\lambda_1)^C P^C_{z_C}(\uparrow|\lambda_1)\frac{N_{C,1}}{N}+P^C_{z_A}(\uparrow|\lambda_2)^C P^C_{z_B}(\uparrow|\lambda_2)^C P^C_{z_C}(\uparrow|\lambda_2)\frac{N-N_{C,1}}{N}. \tag{5.133}$$

It seems there are no more variables available to satisfy these constraints. It may turn out they are satisfied automatically. It could be that the PHV approach only works for a specific pair of observers of this triplet, yet this seems cheating, for which pair of observers would then be so special for which it works? It may be clear that in this direction as well, further work is required.

Chapter 6

Conclusions

The interpretation of Quantum Mechanics is today still a matter of debate. Among other aspects, in particular its allowance of non-local interactions is unacceptable in a deterministic world view. This interaction can be modelled partly in a local deterministic hidden variable theory, but never completely. The Bell inequalities form a proof for this statement. The inequalities impose an upper limit to the expectation values of physical quantities of two particles described by any local hidden variable theory, no matter whether the variable is deterministic or stochastic. Quantum Mechanics violates these inequalities. In the derivation the measurements are done in the hypothetical EPR Bohm set up, but real experiments in equivalent settings have been done showing results favouring Quantum Mechanics. The theory seems therefore irreplaceable by any alternative local theory.

Yet attempts persist, of which the relatively young approach of 't Hooft is an example. His model of Cellular Automata (CA) describes the world as consisting of deterministically behaving cells on the Planck scale. Each cell has some physical degrees of freedom and only affects its direct neighbours through a time reversible evolution law, so the interactions are entirely local. There can also be evolution laws which are not reversible, in the sense that the past does not follow unambiguously from the present and the system suffers from memory loss. The set of all possible configurations form the ontological basis and the system is always in exactly one of the ontological states. Quantum Mechanics, he argues, is only a tool for doing statistics to describe the complicated behaviour of these many cells. The cells are then described in terms of changeables (and their eigenvalues), which do not correspond to the real properties or ontological states of the system. Although the CA model is a simple first attempt, it seems it can roughly reproduce features of a Quantum

Field Theory, (some serious issues put aside).

In the derivation of the Bell inequalities, using the EPR Bohm set up, independence is assumed between the production of a specific pair of two spin particles by a source and Alice's and Bob's choice of axis to measure the spin against. The interpretation with the CA model would consider this assumption inadmissible. As the whole system behaves deterministically, there is always some cause in the past where both the source and Alice (and/or Bob) were affected by. The behaviour of all three actors can therefore not be seen as independent. In a very rough approximation, the number of physical states of the cells can be related to the memory loss of a system if it is defined in a specific way. It turns out the number of physical states needs to be enormous ($\approx 10^{35}$) for a cell to reliably remember events which took place a macroscopic spacetime distance apart. It seems therefore that imposing only little memory loss already imposes problems on CA to which Bell's Inequalities do not apply.

Another young attempt to create an alternative to Quantum Mechanics is the Perspectival Hidden Variable (PHV) approach. With the philosophy that any strange hidden variable is better than no hidden variable, it proposes to introduce one which is frame dependent and stochastic. The frames of an observer is defined by the choice of a certain measurement set up, for the measurement of the observable spin for example, the orientation of axis along which the spin is measured. Any measurement on a bipartite pure state $\in \mathscr{H}_1 \otimes \mathscr{H}_2$ defines two frames of reference, one corresponding to a measurement of an observable of an operator acting in $\mathscr{H}_1$, the other in $\mathscr{H}_2$ (Alice and Bob). As is required for any stochastic hidden variable (if it can be measured in some way), the value the PHV λ specifies a probability to measure a certain physical quantity. Yet the value of this λ is assumed to depend on the frame of the observer. When the two observers are in the same frame, they would measure the same value of λ. A probabilistic transformation law can be defined relating the values of λ from one frame to another. Alice's spin probabilities then only depend on her choice of axis and Bob's on his choice of axis, but the two of them disagree on the values of these untransformed probabilities. The transformed probabilities, so the probability Alice perceives Bob to have, and vice versa, do depend on both their orientations. It is therefore a matter of debate if parameter independence is broken.

In this work it was assumed λ could only take two possible values. It was then investigated whether this PHV approach could describe the EPR Bohm situation. First the transformations were defined, where it was imposed the transformation from Alice to Bob was the inverse of the transformation from Bob to Alice. Then it was demanded that the probabilities calculated by Alice and by Bob, ignoring the information they have about λ, were the same. Only then consistency with Quantum Mechanics can be achieved. This requirement fixed the probabilities in terms of the probability of the source to produce a certain λ. Then it was investigated more specifically if for this two λ case, the right angle dependence of the Quantum Mechanical probabilities could be reproduced. For the singlet and triplet states, the only solutions obtained were those for a limited yet significant range of the angles. It is expected that introducing more than two λs could give better solutions, yet it turns out the equations which fix the transformation in terms of the source probabilities are very tedious to solve for three λs. One can in tri- or higher partite states also introduce more than two observers. Quantum Mechanics imposes them to all agree if the PHV information is ignored. The consistency of this approach should still be investigated for those cases.

Bibliography

[1] Bell, J.S., *On the Einstein Podolsky Rosen Paradox.* Physics 1, 3, 195200 (1964)

[2] Einstein, A., Podolsky, B., Rosen, N., *Can quantum-mechanical description of physical reality be considered complete?* Phys. Rev. 47 777 (1935)

[3] Bohm, D.,Aharonov, Y.,*Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky* Phys.Rev. 108, 1070-1076 (1957)

[4] Hilgevoord, J. *Foundations of Quantum Mechanics* (Lecture notes with the course 'Foundations of Quantum Mechanics' taught at Utrecht University), (2009)

[5] Clauser, J.F.; Horne, M.A.; Shimony, A.; Holt R.A., *Proposed experiment to test local hidden-variable theories* Phys. Rev. Lett. 23 (15): 8804 (1969)

[6] There have been multiple experiments, most famous is the series by Aspect, see for example Aspect, A.;Grangier, P.; Roger, G., *Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities* Phys. Rev. Lett. 49 (2): 914 (1982)

[7] See for example Wimmel, H. *Quantum Physics and observed reality*, World Scientific Publishing (1992) or Faye, J. *Copenhagen Interpretation of Quantum Mechanics*, The Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/archives/fall2008/entries/qm-copenhagen/(2008)

[8] Styer, D., F.; Balkin, M. S.; Becker, K. M.; Burns, M. R.; Dudley, C. E.; Forth, S. T.; Gaumer, J. S.; Kramer, M. A. et al. *Nine formulations of quantum mechanics*, American Journal of Physics 70 (3): 288297.(March 2002)

[9] Hooft, G. 't, *Hilbert space in deterministic theories, a reconsideration of the interpretation of quantum mechanics* in Proceedings of the 3rd Stueckelberg

Workshop on Relativistic Field Theories, N. Carlevaro, R. Ruffini and G.V. Vereshchagin, eds., Cambridge Scientific Publ., pp. 1-18 (2010)

[10] Hooft, G. 't, *The emergence of Quantum Mechanics*, FFP11, Paris. ITP-UU-10/44, Spin-10/37 (2010)

[11] Hooft, G. 't, *Duality Between a Deterministic Cellular Automaton and a Bosonic Quantum Field Theory in 1+1 Dimensions* Foundations of physics, Volume: 43, Issue: 5 , pp: p597, 18p (2013)

[12] Hooft, G. 't, *Quantum Mechanics and determinism* Proceedings of the Eighth International Conference on 'Particles, Strings and Cosmology, pp. 275-285 (2001)

Appendix A

First Appendix

Marginal and Conditional Probabilities in EPR Bohm according to QM

We will state in detail how for a general state the different probabilities are calculated in Quantum Mechanics (QM). We consider the simplest bipartite states there are, that of two spin 1/2 particles of which we only consider the spin, ignoring any other property. We suppose Alice to do a measurement on one particle and Bob on the other. Alice and Bob can change the probabilities of measuring up and down by adjusting the axis they measure the spin against. As stated in chapter 2, we take the spin of a particle in the $\vec{n}$ direction to be represented by the operator

$$\vec{S} = \frac{\hbar}{2}\vec{n} \cdot \vec{\sigma}, \tag{A.1}$$

where $\vec{\sigma}$ is a vector with the Pauli matrices conventionally chosen. We choose for spherical coordinates for $\vec{n}$

$$\vec{n} = \begin{pmatrix} sin\theta cos\phi \\ sin\theta sin\phi \\ cos\theta \end{pmatrix}, \tag{A.2}$$

and therefore

$$\vec{n} \cdot \vec{\sigma} = \begin{pmatrix} cos\theta & e^{-i\phi} sin\theta \\ e^{i\phi} sin\theta & -cos\theta \end{pmatrix}. \tag{A.3}$$

The eigenvectors of this matrix, with respective eigenvalues $+1$ (up) and -1 (down), can be shown to be

$$|\vec{n}\uparrow\rangle = \begin{pmatrix} e^{\frac{-i\phi}{2}} cos\frac{\theta}{2} \\ e^{\frac{i\phi}{2}} sin\frac{\theta}{2} \end{pmatrix}, \qquad |\vec{n}\downarrow\rangle = \begin{pmatrix} -e^{\frac{-i\phi}{2}} sin\frac{\theta}{2} \\ e^{\frac{i\phi}{2}} cos\frac{\theta}{2} \end{pmatrix}. \tag{A.4}$$

$P^J(\uparrow,\uparrow)$ (and thereby P^M and P^C) for a certain choice of axes $\vec{n}$ and $\vec{m}$ is then calculated as

$$P_J(\vec{n}\uparrow,\vec{m}\uparrow) = |(\langle\vec{n}\uparrow| \otimes \langle\vec{m}\uparrow|)\,|\Psi\rangle\,|^2. \tag{A.5}$$

A general bipartite pure state $|\Psi\rangle$ of two spin 1/2 particles can always be written as

$$|\Psi\rangle = a\,|z\uparrow\rangle_A\,|z\uparrow\rangle_B + b\,|z\uparrow\rangle_A\,|z\downarrow\rangle_B + c\,|z\downarrow\rangle_A\,|z\uparrow\rangle_B + d\,|z\downarrow\rangle_A\,|z\downarrow\rangle_B\,, \tag{A.6}$$

for some normalized complex coefficients $|a|^2+|b|^2+|c|^2+|d|^2 = 1$ and $|z\uparrow\rangle_A$ being the spin up eigenstate of Alice's particle in the z-direction, $|z\downarrow\rangle_A$ the spin down eigenstate in the z-direction, and likewise for Bob's particle. Now let us suppose both Alice and Bob are doing spin measurements along the z axis. It follows the marginal probability (marginal) for Alice to measure up is then

$$P^M_{z_A}(\uparrow) = |a|^2 + |b|^2, \tag{A.7}$$

and the marginal for Bob to measure up

$$P^M_{z_B}(\uparrow) = |a|^2 + |c|^2 \tag{A.8}$$

The marginals to measure spin down are then simply obtained by taking $|\delta|^2$ for $|\alpha|^2$ and $|\beta|^2$ for $|\gamma|^2$ and vice versa, or since they are normalized to 1, take 1−the above values. The conditional probability ('conditional') for Alice to measure up given that Bob measures up is

$$P^C_{z_A,z_B}(\uparrow\mid\uparrow) = \frac{|a|^2}{|a|^2+|c|^2}, \tag{A.9}$$

and for Bob given that Alice measures up:

$$P^C_{z_B,z_A}(\uparrow\mid\uparrow) = \frac{|a|^2}{|a|^2+|b|^2}. \tag{A.10}$$

Again, to obtain the conditionals to measure down given that the other party measures down, we just take $|\delta|^2$ for $|\alpha|^2$, and $|\beta|^2$ for $|\gamma|^2$ and vice versa. The conditional to measure up given that the other party measures down then simply follows from taking by taking 1− the conditional to measure down given that the other party measures down. The other options follow similarly.

If one would want to choose the axes differently, this is a matter of simple substitution of:

$$|z\uparrow\rangle \;=\; e^{\frac{i\phi}{2}}\cos\frac{\theta}{2}\,|\vec{m}\uparrow\rangle - e^{\frac{i\phi}{2}}\sin\frac{\theta}{2}\,|\vec{m}\downarrow\rangle\,, \tag{A.11}$$

$$|z\downarrow\rangle \;=\; e^{\frac{-i\phi}{2}}\sin\frac{\theta}{2}\,|\vec{m}\uparrow\rangle + e^{\frac{-i\phi}{2}}\cos\frac{\theta}{2}\,|\vec{m}\downarrow\rangle\,. \tag{A.12}$$

$$|\vec{m}\uparrow\rangle = \cos\frac{\theta_m}{2}e^{-i\frac{\phi_m}{2}}\,|z\uparrow\rangle + \sin\frac{\theta_m}{2}e^{i\frac{\phi_m}{2}}\,|z\downarrow\rangle\,, \tag{A.13}$$

$$|\vec{m}\downarrow\rangle = -\sin\frac{\theta_m}{2}e^{-i\frac{\phi_m}{2}}\,|z\uparrow\rangle + \cos\frac{\theta_m}{2}e^{i\frac{\phi_m}{2}}\,|z\downarrow\rangle\,. \tag{A.14}$$

Let us suppose Bob chooses his axis z' along $\vec{m}(\theta_m, \phi_m)$, so the state becomes in those terms:

$$\begin{aligned} |\Psi\rangle = a'(\theta_m,\phi_m)\,|z\uparrow\rangle_A\,|\vec{m}\uparrow\rangle_B + b'(\theta_m,\phi_m)\,|z\uparrow\rangle_A\,|\vec{m}\downarrow\rangle_B + \\ c'(\theta_m,\phi_m)\,|z\downarrow\rangle_A\,|\vec{m}\uparrow\rangle_B + d'(\theta_m,\phi_m)\,|z\downarrow\rangle_A\,|\vec{m}\downarrow\rangle_B\,, \end{aligned} \tag{A.15}$$

where the coefficients (and hence the joint probabilities) are given by

$$a'(\theta,\phi) = e^{\frac{i\phi}{2}}\cos\frac{\theta}{2}a + e^{\frac{-i\phi}{2}}\sin\frac{\theta}{2}b, \tag{A.16}$$

$$b'(\theta,\phi) = -e^{\frac{i\phi}{2}}\sin\frac{\theta}{2}a + e^{\frac{-i\phi}{2}}\cos\frac{\theta}{2}b, \tag{A.17}$$

$$c'(\theta,\phi) = e^{\frac{i\phi}{2}}\cos\frac{\theta}{2}c + e^{\frac{-i\phi}{2}}\sin\frac{\theta}{2}d, \tag{A.18}$$

$$d'(\theta,\phi) = -e^{\frac{i\phi}{2}}\sin\frac{\theta}{2}c + e^{\frac{-i\phi}{2}}\cos\frac{\theta}{2}d. \tag{A.19}$$

We are still considering how QM then tells us how the probabilities are calculated given such a state. Alice measures a marginal probability for spin up given by

$$P^M_{z_A}(\uparrow) = |a'|^2 + |b'|^2 = |a|^2 + |b|^2, \tag{A.20}$$

and Bob measures a marginal for spin up

$$\begin{aligned} &P^M_{z_B}(\uparrow) = |a'|^2 + |c'|^2 = \\ &\cos^2\frac{\theta_m}{2}(|a|^2+|c|^2)+\sin^2\frac{\theta_m}{2}(|b|^2+|d|^2)+\frac{1}{2}\sin\theta_m(ab^*e^{i\phi_m}+ba^*e^{-i\phi_m}+cd^*e^{i\phi_m}+dc^*e^{-i\phi_m}). \end{aligned} \tag{A.21}$$

The conditionals can be calculated from these and the expressions for the joint probabilities. We see that for the case of the singlet, $|a| = |d| = 0$ and $|c| = |b| = 1/\sqrt{2}$, we get

$$|a'(\theta,\phi)|^2 = \frac{1}{2}\sin^2\frac{\theta}{2}, \tag{A.22}$$
$$|b'(\theta,\phi)|^2 = \frac{1}{2}\cos^2\frac{\theta}{2}, \tag{A.23}$$
$$|c'(\theta,\phi)|^2 = \frac{1}{2}\cos^2\frac{\theta}{2}, \tag{A.24}$$
$$|d'(\theta,\phi)|^2 = \frac{1}{2}\sin^2\frac{\theta}{2}. \tag{A.25}$$

So that the marginals are both $1/2$ and the conditionals for Alice and Bob for opposite spin $\cos^2\frac{\theta}{2}$ and for similar spin $\sin^2\frac{\theta}{2}$. Next we will consider what the expressions for the coefficients will be when both Alice and Bob choose their axes different from those in which the state is originally formulated. In that way we can separate the dependence on Bob's and Alice's axes. Inserting the expressions (A.11) for both Alice's and Bob's kets, and rearranging, the following expressions are obtained.

$$\begin{aligned}\alpha(\theta_1,\phi_1,\theta_2,\phi_2) &= e^{\frac{i(\phi_1+\phi_2)}{2}}cos\frac{\theta_1}{2}cos\frac{\theta_2}{2}a + e^{\frac{i(\phi_1-\phi_2)}{2}}cos\frac{\theta_1}{2}sin\frac{\theta_2}{2}b \\ &+e^{\frac{i(\phi_2-\phi_1)}{2}}sin\frac{\theta_1}{2}cos\frac{\theta_2}{2}c + e^{-\frac{i(\phi_1+\phi_2)}{2}}sin\frac{\theta_1}{2}sin\frac{\theta_2}{2}d, \quad \text{(A.26)}\\ \beta(\theta_1,\phi_1,\theta_2,\phi_2) &= -e^{\frac{i(\phi_1+\phi_2)}{2}}cos\frac{\theta_1}{2}sin\frac{\theta_2}{2}a + e^{\frac{i(\phi_1-\phi_2)}{2}}cos\frac{\theta_1}{2}cos\frac{\theta_2}{2}b \\ &-e^{\frac{i(\phi_2-\phi_1)}{2}}sin\frac{\theta_1}{2}sin\frac{\theta_2}{2}c + e^{-\frac{i(\phi_1+\phi_2)}{2}}sin\frac{\theta_1}{2}cos\frac{\theta_2}{2}d, \quad \text{(A.27)}\\ \gamma(\theta_1,\phi_1,\theta_2,\phi_2) &= -e^{\frac{i(\phi_1+\phi_2)}{2}}sin\frac{\theta_1}{2}cos\frac{\theta_2}{2}a - e^{\frac{i(\phi_1-\phi_2)}{2}}sin\frac{\theta_1}{2}sin\frac{\theta_2}{2}b \\ &+e^{\frac{i(\phi_2-\phi_1)}{2}}cos\frac{\theta_1}{2}cos\frac{\theta_2}{2}c + e^{-\frac{i(\phi_1+\phi_2)}{2}}cos\frac{\theta_1}{2}sin\frac{\theta_2}{2}d, \quad \text{(A.28)}\\ \delta(\theta_1,\phi_1,\theta_2,\phi_2) &= e^{\frac{i(\phi_1+\phi_2)}{2}}sin\frac{\theta_1}{2}sin\frac{\theta_2}{2}a - e^{\frac{i(\phi_1-\phi_2)}{2}}sin\frac{\theta_1}{2}cos\frac{\theta_2}{2}b \\ &-e^{\frac{i(\phi_2-\phi_1)}{2}}cos\frac{\theta_1}{2}sin\frac{\theta_2}{2}c + e^{-\frac{i(\phi_1+\phi_2)}{2}}cos\frac{\theta_1}{2}cos\frac{\theta_2}{2}d. \quad \text{(A.29)}\end{aligned}$$

Appendix B

Second Appendix

Appendix 2 Checking inverse probability transformations

We will check that for 3 λs the probability transformation nxn matrix for Bob is the inverse of Alice's. It is then fairly easy to see how it generalizes for any number of λs. For three λs we have the transformation rule of Bob's conditional probabilities to Alice's:

$$\begin{pmatrix} P_{z_B}(\uparrow|\lambda_1)^A \\ P_{z_B}(\uparrow|\lambda_2)^A \\ P_{z_B}(\uparrow|\lambda_3)^A \end{pmatrix} =$$

$$\begin{pmatrix} (1-p_{1,2}-p_{1,3})\frac{N_{B,1}}{N^A_{B,1}} & p_{2,1}\frac{N_{B,2}}{N^A_{B,1}} & p_{3,1}\frac{N_{B,3}}{N^A_{B,1}} \\ p_{1,2}\frac{N_{B,1}}{N^A_{B,2}} & (1-p_{2,1}-p_{2,3})\frac{N_{B,2}}{N^A_{B,2}} & p_{3,2}\frac{N_{B,3}}{N^A_{B,2}} \\ p_{1,3}\frac{N_{B,1}}{N^A_{B,3}} & p_{2,3}\frac{N_{B,2}}{N^A_{B,3}} & (1-p_{3,1}-p_{3,2})\frac{N_{B,3}}{N^A_{B,3}} \end{pmatrix} \begin{pmatrix} P_{z_B}(\uparrow|\lambda_1) \\ P_{z_B}(\uparrow|\lambda_2) \\ P_{z_B}(\uparrow|\lambda_3) \end{pmatrix}.$$

We ought to show that if we make the substitutions:

$$\begin{pmatrix} P_{z_B}(\uparrow|\lambda_1)^A \\ P_{z_B}(\uparrow|\lambda_2)^A \\ P_{z_B}(\uparrow|\lambda_3)^A \end{pmatrix} \to \begin{pmatrix} P_{z_A}(\uparrow|\lambda_1)^B \\ P_{z_A}(\uparrow|\lambda_2)^B \\ P_{z_A}(\uparrow|\lambda_3)^B \end{pmatrix},$$

$$\begin{pmatrix} P_{z_B}(\uparrow|\lambda_1) \\ P_{z_B}(\uparrow|\lambda_2) \\ P_{z_B}(\uparrow|\lambda_3) \end{pmatrix} \to \begin{pmatrix} P_{z_A}(\uparrow|\lambda_1) \\ P_{z_A}(\uparrow|\lambda_2) \\ P_{z_A}(\uparrow|\lambda_3) \end{pmatrix},$$

equality arises if the matrix M_{AB} is substituted for its inverse M_{AB}^{-1}. We start with rewriting the starting expression as

$$\begin{pmatrix} P_{z_B}(\uparrow|\lambda_1)^A \\ P_{z_B}(\uparrow|\lambda_2)^A \\ P_{z_B}(\uparrow|\lambda_3)^A \end{pmatrix} = \begin{pmatrix} N_{B,1} & & \\ & N_{B,2} & \\ & & N_{B,3} \end{pmatrix} \cdot$$

$$\begin{pmatrix} 1-p_{1,2}-p_{1,3} & p_{2,1} & p_{3,1} \\ p_{1,2} & 1-p_{2,1}-p_{2,3} & p_{3,2} \\ p_{1,3} & p_{2,3} & 1-p_{3,1}-p_{3,2} \end{pmatrix} \begin{pmatrix} \frac{1}{N_{B,1}^A} & & \\ & \frac{1}{N_{B,2}^A} & \\ & & \frac{1}{N_{B,3}^A} \end{pmatrix} \begin{pmatrix} P_{z_B}(\uparrow|\lambda_1) \\ P_{z_B}(\uparrow|\lambda_2) \\ P_{z_B}(\uparrow|\lambda_3) \end{pmatrix}.$$

Now we substitute the matrix for its inverse, where P^{-1} denotes the inverse of the middle matrix,

$$\begin{pmatrix} P_{z_B}(\uparrow|\lambda_1^A) \\ P_{z_B}(\uparrow|\lambda_2^A) \\ P_{z_B}(\uparrow|\lambda_3^A) \end{pmatrix} = \begin{pmatrix} \frac{1}{N_{B,1}} & & \\ & \frac{1}{N_{B,2}} & \\ & & \frac{1}{N_{B,3}} \end{pmatrix} P^{-1} \begin{pmatrix} N_{B,1}^A & & \\ & N_{B,2}^A & \\ & & N_{B,3}^A \end{pmatrix} \begin{pmatrix} P_{z_B}(\uparrow|\lambda_1) \\ P_{z_B}(\uparrow|\lambda_2) \\ P_{z_B}(\uparrow|\lambda_3) \end{pmatrix}.$$

Now we postulate

$$N_{B,1}^A = N_{A,1}, \qquad N_{B,2}^A = N_{A,2}, \qquad N_{B,3}^A = N_{A,3}, \tag{B.1}$$

which can be shown to be equivalent to

$$N_{B,1} = N_{A,1}^B, \qquad N_{B,2} = N_{A,2}^B, \qquad N_{B,3} = N_{A,3}^B. \tag{B.2}$$

Next we substitute these and we also make the substitution for the probability vectors

$$\begin{pmatrix} P_{z_A}(\uparrow|\lambda_1)^B \\ P_{z_A}(\uparrow|\lambda_2)^B \\ P_{z_A}(\uparrow|\lambda_3)^B \end{pmatrix} = \begin{pmatrix} \frac{1}{N_{A,1}^B} & & \\ & \frac{1}{N_{A,2}^B} & \\ & & \frac{1}{N_{A,3}^B} \end{pmatrix} P^{-1} \begin{pmatrix} N_{A,1} & & \\ & N_{A,2} & \\ & & N_{A,3} \end{pmatrix} \begin{pmatrix} P_{z_A}(\uparrow|\lambda_1) \\ P_{z_A}(\uparrow|\lambda_2) \\ P_{z_A}(\uparrow|\lambda_3) \end{pmatrix},$$

$$\begin{pmatrix} n_{\uparrow,A,1}^B/N_{A,1}^B \\ n_{\uparrow,A,2}^B/N_{A,2}^B \\ n_{\uparrow,A,3}^B/N_{A,3}^B \end{pmatrix} = \begin{pmatrix} \frac{1}{N_{A,1}^B} & & \\ & \frac{1}{N_{A,2}^B} & \\ & & \frac{1}{N_{A,3}^B} \end{pmatrix} P^{-1} \begin{pmatrix} N_{A,1} & & \\ & N_{A,2} & \\ & & N_{A,3} \end{pmatrix} \begin{pmatrix} n_{\uparrow,A,1}/N_{A,1} \\ n_{\uparrow,A,2}/N_{A,2} \\ n_{\uparrow,A,3}/N_{A,2} \end{pmatrix},$$

$$\begin{pmatrix} n_{\uparrow,A,1}^B \\ n_{\uparrow,A,2}^B \\ n_{\uparrow,A,3}^B \end{pmatrix} = P^{-1} \begin{pmatrix} n_{\uparrow,A,1} \\ n_{\uparrow,A,2} \\ n_{\uparrow,A,3} \end{pmatrix}.$$

This is true as P^{-1} is defined to be the transformation matrix of the vector with the populations $n_{\uparrow,A,n}$. It is not hard to see how this generalized to matrices of any dimension.

Appendix C

Third Appendix

Going from (5.69) to (5.71)

In this appendix we will simplify the algebra of (5.69) to (5.71). We start one step before from

$$P^{QM}_{z_B,z_A}(\uparrow,\uparrow) = \frac{1}{2N}$$
$$\Big(P^C_{z_B}(\uparrow|\lambda_1)\big(\frac{8\frac{N^2_{B,1}}{N} - 8\frac{N^3_{B,1}}{N^2} - 2N_{B,1}}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1) + 4(N_{B,1} - \frac{N^2_{B,1}}{N})P^C_{z_A}(\uparrow|\lambda_2)\big)$$
$$+P^C_{z_B}(\uparrow|\lambda_2)\big(4(N_{B,1} - \frac{N^2_{B,1}}{N})P^C_{z_A}(\uparrow|\lambda_2) + \frac{16\frac{N^2_{B,1}}{N} - 8\frac{N^3_{B,1}}{N^2} - 10N_{B,1} + 2N}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1)\big)\Big) \tag{C.1}$$

$$= P^C_{z_B}(\uparrow|\lambda_1)\big(\frac{4\frac{N^2_{B,1}}{N^2} - 4\frac{N^3_{B,1}}{N^3} - \frac{N_{B,1}}{N}}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1) + 2(\frac{N_{B,1}}{N} - \frac{N^2_{B,1}}{N^2})P^C_{z_A}(\uparrow|\lambda_2)\big)$$
$$+ P^C_{z_B}(\uparrow|\lambda_2)\big(2(\frac{N_{B,1}}{N} - \frac{N^2_{B,1}}{N^2})P^C_{z_A}(\uparrow|\lambda_2) + \frac{8\frac{N^2_{B,1}}{N^2} - 4\frac{N^3_{B,1}}{N^3} - 5\frac{N_{B,1}}{N} + 1}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1)\big) \tag{C.2}$$

Now substituting the expressions resulting from the calculations of the marginal probabilities (5.62) and (5.65)

$$
\begin{aligned}
&P^{QM}_{z_B,z_A}(\uparrow,\uparrow) = P^C_{z_B}(\uparrow|\lambda_1)\cdot \\
&\Big(\frac{4\frac{N^2_{B,1}}{N^2}-4\frac{N^3_{B,1}}{N^3}-\frac{N_{B,1}}{N}}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1)+2(\frac{N_{B,1}}{N}-\frac{N^2_{B,1}}{N^2})\frac{P^C_{z_A}(\uparrow|\lambda_{A,1})(\frac{N^2_{B,1}}{N^2}-\frac{N_{B,1}}{N}+\frac{1}{4})-\frac{|\alpha|^2+|\beta|^2}{4}}{\frac{N^2_{B,1}}{N^2}-\frac{N_{B,1}}{N}}\Big) \\
&\qquad + \frac{P^C_{z_B}(\uparrow|\lambda_1)N_{B,1}/N-(|\alpha|^2+|\gamma|^2)}{N_{B,1}/N-1}\cdot \\
&\big(2(\frac{N_{B,1}}{N}-\frac{N^2_{B,1}}{N^2})\frac{P^C_{z_A}(\uparrow|\lambda_1)(\frac{N^2_{B,1}}{N^2}-\frac{N_{B,1}}{N}+\frac{1}{4})-\frac{|\alpha|^2+|\beta|^2}{4}}{\frac{N^2_{B,1}}{N^2}-\frac{N_{B,1}}{N}}+\frac{8\frac{N^2_{B,1}}{N^2}-4\frac{N^3_{B,1}}{N^3}-5\frac{N_{B,1}}{N}+1}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1)\big)
\end{aligned}
\tag{C.3}
$$

$$
\begin{aligned}
&= P^C_{z_B}(\uparrow|\lambda_1)\big((\frac{4\frac{N^2_{B,1}}{N^2}-4\frac{N^3_{B,1}}{N^3}-\frac{N_{B,1}}{N}}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1)+P^C_{z_A}(\uparrow|\lambda_{A,1})(-\frac{2N^2_{B,1}}{N^2}+2\frac{N_{B,1}}{N}-\frac{1}{2})+\frac{|\alpha|^2+|\beta|^2}{2}) \\
&\qquad + \frac{P^C_{z_B}(\uparrow|\lambda_1)N_{B,1}/N-(|\alpha|^2+|\gamma|^2)}{N_{B,1}/N-1}\cdot \\
&\big(P^C_{z_A}(\uparrow|\lambda_1)(-2\frac{N^2_{B,1}}{N^2}+2\frac{N_{B,1}}{N}-\frac{1}{2})+\frac{|\alpha|^2+|\beta|^2}{2}+\frac{8\frac{N^2_{B,1}}{N^2}-4\frac{N^3_{B,1}}{N^3}-5\frac{N_{B,1}}{N}+1}{1-2\frac{N_{B,1}}{N}}P^C_{z_A}(\uparrow|\lambda_1)\big).
\end{aligned}
\tag{C.4}
$$

In the following we will write $N_{B,1}/N = x$ (we could write $q_{B,1} = x$ taking the large N limit):

$$
\begin{aligned}
P^{QM}_{z_B,z_A}(\uparrow,\uparrow) &= P^C_{z_B}(\uparrow|\lambda_1)(\frac{-2x^2+2x-1/2}{1-2x}P^C_{z_A}(\uparrow|\lambda_1)+\frac{|\alpha|^2+|\beta|^2}{2}) \\
&\quad + \frac{P^C_{z_B}(\uparrow|\lambda_1)x-|\alpha|^2+|\gamma|^2}{x-1}(P^C_{z_A}(\uparrow|\lambda_1)\frac{2x^2-2x+1/2}{1-2x}+\frac{|\alpha|^2+|\beta|^2}{2})
\end{aligned}
\tag{C.5}
$$

$$
\begin{aligned}
&= P^C_{z_B}(\uparrow|\lambda_1)((x-\frac{1}{2})P^C_{z_A}(\uparrow|\lambda_1)+\frac{|\alpha|^2+|\beta|^2}{2}) \\
&\qquad + \frac{P^C_{z_B}(\uparrow|\lambda_1)x-(|\alpha|^2+|\gamma|^2)}{x-1}(P^C_{z_A}(\uparrow|\lambda_1)(\frac{1}{2}-x)+\frac{|\alpha|^2+|\beta|^2}{2})
\end{aligned}
\tag{C.6}
$$

$$= P^C_{z_B}(\uparrow|\lambda_1)(x-\frac{1}{2})P^C_{z_A}(\uparrow|\lambda_1) + P^C_{z_B}(\uparrow|\lambda_1)\frac{|\alpha|^2+|\beta|^2}{2}$$
$$+\frac{P^C_{z_B}(\uparrow|\lambda_1)x-(|\alpha|^2+|\gamma|^2)}{x-1}P^C_{z_A}(\uparrow|\lambda_1)(\frac{1}{2}-x)+\frac{|\alpha|^2+|\beta|^2}{2}\frac{P^C_{z_B}(\uparrow|\lambda_1)x-(|\alpha|^2+|\gamma|^2)}{x-1} \tag{C.7}$$

$$= \frac{P^C_{z_B}(\uparrow|\lambda_1)P^C_{z_A}(\uparrow|\lambda_1)(x-\frac{1}{2})(x-1)+(P^C_{z_B}(\uparrow|\lambda_1)x-(|\alpha|^2+|\gamma|^2))P^C_{z_A}(\uparrow|\lambda_1)(\frac{1}{2}-x)}{x-1}$$
$$+\frac{P^C_{z_B}(\uparrow|\lambda_1)x-(|\alpha|^2+|\gamma|^2)+P^C_{z_B}(\uparrow|\lambda_1)(x-1)}{x-1}\frac{|\alpha|^2+|\beta|^2}{2} \tag{C.8}$$

$$= \frac{P^C_{z_B}(\uparrow|\lambda_1)P^C_{z_A}(\uparrow|\lambda_1)(\frac{1}{2}-x)-(|\alpha|^2+|\gamma|^2)P^C_{z_A}(\uparrow|\lambda_{A,1})(\frac{1}{2}-x)}{x-1}$$
$$+\frac{-(|\alpha|^2+|\gamma|^2)+P^C_{z_B}(\uparrow|\lambda_1)(2x-1)}{x-1}\frac{|\alpha|^2+|\beta|^2}{2} \tag{C.9}$$

$$= \frac{(\frac{1}{2}-x)(P^C_{z_B}(\uparrow|\lambda_1)P^C_{z_A}(\uparrow|\lambda_1)-(|\alpha|^2+|\gamma|^2)P^C_{z_A}(\uparrow|\lambda_{A,1})-(|\alpha|^2+|\beta|^2)P^C_{z_B}(\uparrow|\lambda_{A,1}))}{x-1}$$
$$-\frac{(|\alpha|^2+|\gamma|^2)(|\alpha|^2+|\beta|^2)}{2(x-1)} \tag{C.10}$$

$$= \frac{(\frac{1}{2}-x)(P^C_{z_B}(\uparrow|\lambda_1)-(|\alpha|^2+|\gamma|^2))(P^C_{z_A}(\uparrow|\lambda_1)-(|\alpha|^2+|\beta|^2))}{x-1}$$
$$-\frac{(\frac{1}{2}-x)(|\alpha|^2+|\beta|^2)(|\alpha|^2+|\gamma|^2)}{x-1}-\frac{(|\alpha|^2+|\gamma|^2)(|\alpha|^2+|\beta|^2)}{2(x-1)} \tag{C.11}$$

$$P^{QM}_{z_B,z_A}(\uparrow,\uparrow) = \frac{(\frac{1}{2}-x)(P^C_{z_B}(\uparrow|\lambda_1)-(|\alpha|^2+|\gamma|^2))(P^C_{z_A}(\uparrow|\lambda_1)-(|\alpha|^2+|\beta|^2))}{x-1}$$
$$+(|\alpha|^2+|\beta|^2)(|\alpha|^2+|\gamma|^2). \tag{C.12}$$

This is the desired form.