Seminar
Foundations of Quantum Mechanics

Organiser: Robert C. Helling
Ludwig-Maximilians-Universität München

Winter semester 2010/11

Contents

Preface

Even in the twenty-first century, there are discussion on the foundations of quantum mechanics many of them so heated that most physicists try to stay away from them. Nevertheless, Kostas Vavouranakis and Kosmas Kepesidis convinced me that as part of the master program "Theoretical and Mathematical Physics" we should have a seminar on this topic which was realized in the winter semester 2010/11. My condition for running this seminar was that it should be at a "no nonsense" level and that I would stop any violation immediately. Luckily I did not have to do this too often. In fact, I was quite impressed by the material assembled by the participants. So, here, I have collected the write-ups of the presentations in book form. My editing has been minimal, mainly typesetting so individual style of the contributors is still clearly visible. I would like to thank all participants of the foundations seminar and hope that they enjoyed it and benefited from it as much as I did!

1. Introduction - Robert C. Helling

1.1. Why this seminar?

Quantum physics enjoys the reputation of being mysterious and not understandable which is perpetuated by popular texts. When taking a lecture course on the subject, however, there appears to be no mystery at all: For a given system, the harmonic oscillator say, one immediately writes down the Hamilton operator as a differential operator and computes its spectrum and eigenfunctions which tells one everything about the problem at hand. Diagonalizing the operator can be difficult in practice but this appears to be merely a technical problem due to the fact that solving partial differential equations is hard, there is no "mystery" associated to it.

So where does quantum mechanics get this reputation of being mysterious? Is it only a historical misunderstanding that came about when our forefathers, in particular Einstein, struggled with the formalism that was only being invented at the time and which lead to a philosophical overloading of the theory in connection with wrong prejudices about how the world should be? "Shut up and calculate" is the approach (wrongly[17]attributed to Feynman) to address (or better not to address) these worries which appear to be only psychological and which seem to not appear or matter in real problems. So why waste our time with a seminar on such non-questions?

There are a few points that the "just calculate the spectrum of the operator" approach does not address and which we found worthwhile to understand a bit better for ourselves: First of all, why does the formulation of quantum physics with a Hilbert space and Hamilton operator work the way it does? Would there have been alternatives? What is forced on us by mathematics starting from some simple assumptions? The other big complex of questions (and probably the origin of much of the confusion and mystery) is concerned with the observation that our day to day macroscopic world appears to be governed by the laws of classical physics and not by the equations of quantum mechanics which do describe the world at small scales. But where is the transition, how do the quantum and the classical world meet and interact? The 'measurement problem' and 'Schrödingers cat' are prominent examples of this set of questions. How can the quantum world of superpositions and probabilities lead to our macroscopic observations in which measurements

do have a definite outcome? One resolution is that maybe the whole world is in fact quantum mechanical, not only on small but on all scales and the classical appearance is only an illusion or at best an approximation. We will see how this classicalness comes about.

Anything titled "Foundations of Quantum Mechanics" is likely to be a discussion in philosophical terms, often with ill-defined concepts. When agreeing to run this seminar, I insisted on a strict "no-nonsense" policy and I believe we succeeded quite well in this respect. Part of which I would like to attribute to the fact that we interpreted "Foundations" at least as often to be "mathematical" rather than "philosophical".

1.2. The Classical and the Quantum

1.2.1. Is the quantum world classical?

Before we go in the direction of explaining the classical world form a quantum perspective, let us first entertain the possibility that it is in fact the other way around: Maybe, deep in its heart, the quantum world is in fact classical. It is just the fact that we have not seen all the relevant fine structure to realize this. After all, in kinetic gas theory, one also deals with a probabilistic theory that explains everything about how gases in a steam engine work but probabilities and their associated vagueness only come about because we ignore the detailed state of all the 10^{23} molecules that in their bouncing around to make up the gas. If we measured all positions and momenta of the molecules, the container of gas would be completely deterministic and the statistical nature is only due to our ignorance, it is not fundamental. Maybe the probabilities in quantum mechanics arise in a similar way.

This possibility, however, is ruled out by Bell-type measurements that maximally violate the associated inequalities, at least if one further assumes locality, the principle that there is no action at a distance. As long as we are dealing with non-relativistic theories this is not fundamental but it seems to hold in macroscopic physics: At least, we have learned to encode forces, that like the gravitational force or the electromagnetic force seem to act between separated systems in terms of fields. These fields then obey local field (differential) equations. This seems to be just a trick, replacing the non- local Coulomb law by the local Gauss-law (and the other Maxwell equations). But

there is more to it: The field is itself a degree of freedom and has dynamics of its own, the electromagnetic waves. Those come about when taking the field as a real physical entity. Of course, one can still generalize the Coulomb force to Liénard-Wiechert potentials but those are then again local in the sense that they depend on the retarded positions and the effect of moving charges around can only effect other charges at the speed of light and not instantly.

The impossibility of having a classical, local theory is most easily seen in the GHZ- M-experiment[12, 15] as presented by Coleman[6]. Consider the following experimental set-up: Three students are each given a "detector" being a box with a port and two buttons labeled 'A' and 'B' and two lights, a red one and a green one. They take their detectors and travel very far in different directions (so that we can assume they are causally disconnected). Each day, they all receive a small package from a central entity (the mothership, professor, etc.) which they plug into the port and then at their liberty press one of the buttons which makes one of the light go on. They take notes in their lab journal of which button they pressed which day and which of the lights turned on as an effect.

Neither we nor the students have any clue about the inner workings of the detectors or the packages. One might speculate that depending on the choice of button the detector compares either the temperature of the package to a reference or measures the weight of the package and shows green if it is lighter than a kilogram. Or the package is empty or completely ignored by the detector which is just a complicated coin throwing device. Or it does a complicated mechanical measurement involving a very complicated system of thousands of gear-wheels. Or the package contains some atoms in an atomic trap and the device does some quantum measurement on them. But if quantum mechanics is classical at it heart, the last choice would again be some complicated classical measurement just in terms of the classical fine structure instead of gear-wheels. We do not know and in fact it does not matter.

After many years of these measurements, the three students meet again and compare their lab journals. They make two peculiar observations: First, on days, on which all three happened to press the 'A' button, the total number of green lights showing was always odd and thus the number of red lights even. On days, however, on which only one of the students pressed 'A' and

the other two pressed 'B' it was the other way around, an even number of green and an odd number of red lights showing. They could not find any correlation between who pressed which button and the light of that person himself, only with the total number.

Can this experiment be explained in terms of classical, local physics? In that case, we could say the detector, depending on the choice of button, measures either the 'A' aspect or the 'B' aspect of the package (which itself could be random). This makes implicitly use of locality: What is measured by student one only depends on his package and his choice of button. It does not depend on which button is chosen by the other two students. Note that we do not make any assumption on the packages and what they contain. Since they come from some unknown preparation it could well be that the contents of the three packages are correlated, for example that the contents is random between days but on each day all three packages have the same content.

To formalize, if the red light shows denote the value of this aspect to be +1 and −1 for the green light. If for example on a particular day student one presses the 'A' button and sees a green light we denote this by $a_1 = -1$ while student three who pressed 'B' sees red which we write as $b_3 = +1$. In this notation, the two observations of the students read

$$a_1 a_2 a_3 = -1 \quad a_1 b_2 b_3 = b_1 a_2 b_3 = b_1 b_2 a_3 = 1$$

Since the only possible values can be ±1 we can express the second observation also as

$$a_1 = b_2 b_3 \quad a_2 = b_1 b_3 \quad a_3 = b_1 b_2$$

But multiplying these three equations leads to a contradiction with the first observations, since then

$$a_1 a_2 a_3 = b_2 b_3 b_1 b_3 b_1 b_2 = 1 \neq -1$$

This seems to be impossible. But there is a simple quantum experiment which yields exactly this outcome: The packages are prepared to each contain a spin 1/2 system that is set-up such that the total wave function is

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\uparrow\rangle - |\downarrow\downarrow\downarrow\rangle)$$

Then, depending on the button, the detector measures either

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{or} \quad \sigma_y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \tag{1.1}$$

The state $|\psi\rangle$ is obviously a -1 eigenstate of three 'A' measurements $\sigma_x \otimes \sigma_x \otimes \sigma_x$ since that operator exchanges the two terms. Swapping two of the σ_x of σ_y still exchanges the two terms but the factor $i^2 = -1$ makes the eigenvalue +1 while two minus-signs cancel. Thus, $|\psi\rangle$ is also a +1 eigenvalue of $\sigma_x \otimes \sigma_y \otimes \sigma_y$ or permutations and the measurements come out as observed by the students. The contradiction of the classical analysis does not arise since quantum mechanically, if the σ_x component of a spin is measured it does not make sense to assign a value to σ_y as it is done in the above calculation. Thus the terms are not defined unless measured but on each day either only a_1 or b_1 has a value. The spin components are not 'realistic', they do not have a value unless observed.

We conclude that using quantum mechanics it is possible to build devices (like the packages and detectors in the example) that cannot be built in an all classical world, at least not unless one allows action at a distance such that the measurement of student one can depend on the choice of button at the detector of students two and three. Thus quantum physics cannot secretly be a classical, local theory.

This tension between realism, contextuality (the requirement that the outcome of measuring a property A should not depend on whether simultaneously one also measures the compatible properties B or C) and locality in the quantum theory is the subject of the chapter on the EPR paradox, Bell type inequalities, and the Kochen-Specker theorem by Isabel Krebs and Matthias Schlaffer.

1.2.2. Are there distinct classical and quantum worlds....

Many textbooks of quantum physics would at least implicitly answer this question affirmatively. This is at the heart of the traditional understanding of quantum mechanics, the "Copenhagen interpretation": One starts with a quantum system that follows some unitary time evolution of the wave function $|\psi\rangle$ governed by the Schrödinger equation. Then, one decides to measure an observable represented by a hermitian operator A. The outcome

is one of the eigenvalues a_i of the operator (with probability $\langle\psi| P_i |\psi\rangle$ where P_i is the projector to the a_i-eigen-space). This measurement makes the wave function collapse such that after the measurement it is the projected wave function

$$\frac{P_i |\psi\rangle}{\sqrt{\langle\psi| P_i |\psi\rangle}} \tag{1.2}$$

such that subsequent measurements of A always yield the same result a_i unless some other observable that does not commute with A is measured intermediately.

This traditional view has obvious short-comings: There are two different, conflicting time-evolutions of quantum states, the unitary one for unobserved states and the col- lapsing one during measurement. And what operation does constitute a measurement that yields a definite, classical outcome for a quantum state? Is it the detector with its macroscopic size or does it need a human brain that reads out (and hopefully understands) the result of the measurement? What about the computer that stores the outcome before some real person looks at it? Apparently, somewhere there should be some sort of transition from the quantum world with its typical superpositions to a classical world where measurements have definite outcomes.

1.2.3. ....or is the classical world secretly quantum?

Reflecting on the differences between the appearance of a classical world and a quantum world one finds the differences are not as obvious as one might have thought. We have seen already above that from statistical mechanics we are used to deal with probabilistic states also in the classical theory by generalizing the notion of a state from points in phase space to probability measures on phase space (of which δ-function supported measures are just special cases). It is perfectly fine to describe the state of a coin after throwing it to be 50% heads and 50% tails. That does not mean that looking at the coin one sees a blurry image but it means that statistically half of the times the outcome is heads and the other half tails and stays with that outcome until the coin is thrown again. (One could have issues with this frequentist notion of probability but the reader can replace this with his favorite interpretation of probability like the fair price one would be willing to pay to play a game that pays one dollar for heads and nothing for tails). This is the same as in quantum mechanics where a wave function with a wide

spread does not describe a thick, squashed electron but an electron whose position measurement can yield a wide variety of results (note well that I did not write "an electron whose unknown position has a wide rage of possible values" as the notion of position of the electron will in general not be defined unless measured). The fact one is dealing with a probabilistic theory is thus not distinguishing quantum mechanics from classical mechanics. One might argue that in the quantum theory, the probabilistic nature is fundamental and in the classical theory is just arises as one did not care to measure the precise state. But to me it seems this distinction between "fundamental" and "effective" does not carry very far when defined operationally.

The next apparent difference between classical physics and quantum mechanics is that in the quantum case, one can take arbitrary superpositions of Hilbert space vectors that represent states (more on that relation below) exemplified in Schrödinger's cat that is in a state which is a superposition of alive and dead. One is tempted to say that one has never seen such a superposition in the macroscopic world. But on second thoughts it is not really clear what such a macroscopic superposition would look like: Ideas about semi-transparent cross-fades of two cats that I might see looking at both states are not warranted even by quantum mechanics, they also do not occur in the microscopic realm. In a double-slit there are not two half, semi-transparent electrons going through each slit, rather whenever you bother to check, you detect an electron either in slit one or slit two. The same holds for the cat: Either there is no interaction between the superimposed cat and your eyes or your eyes after looking at the cat will also be in a superimposed state of either seeing a dead or seeing a living cat. The part state of your eyes that see a dead cat see no trace of a living cat. You never see both states at once. So the conclusion of this discussion has to be that a superposition looks exactly like the states it is composed of, it is just that the looking observer is then also in a superposition.

So then, if it is not the statistical nature and not a nebulous vagueness of super- imposed quantum states that amount to the difference between the classical and the quantum world what is it? This leaves the possibility to have interference in quantum theory, the fact that the wave function has a phase that makes it possible to get less output from opening up more possibilities. This is what the double slit experiment demonstrates. If ψ_1 and ψ_2 are the wave functions describing the particle going through slit one or two, then

the classically available information are the intensities $|\psi_1(x)|^2$ and $|\psi_2(x)|^2$ which would classically add once both slits are open $|\psi_1(x)|^2+|\psi_2(x)|^2$. Here, x is the position on the screen. Classically, if a result can arise from two mutually exclusive possibilities then the probability of one or the other occurring is the sum of the individual probabilities. In the quantum theory, however, it is the amplitudes that add and the resulting probabilities are

$$|psi_1(x) + psi_2(x)|^2 = |\psi_1(x)|^2 + |\psi_2(x)|^2 + 2\mathrm{Re}(\bar{\psi}_1(x)\psi_2(x))$$

There are interference fringes with no particle arriving on the screen at places where particles do arrive when only one of the slits is open: With both slits open, there is destructive interference. It is this effect that in the end makes it possible to build the above described quantum experiment that is classically impossible. The interference term $2\mathrm{Re}(\bar{\psi}_1(x)\psi_2(x))$ depends crucially on the relative phase between ψ_1 and ψ_2. If that is disturbed (for example by interaction with the environment as we will see below) then we come back to the classical addition of probabilities: Let us assume we try to observe if the particle went through slit two. Then ψ_2 has to interact with some sort of detector. There will be a back-reaction and if the rate of particles going through that slit is not to be disturbed, the back-reaction can only be a phase

$$\tilde{\psi}_2(x) = e^{i\varphi}\psi_2(x)$$

But of course φ is not observed and we have to express our lack of knowledge by averaging over it. This results in

$$\frac{1}{2\pi}\int d\varphi\, \mathrm{Re}(\bar{\psi}_1(x)e^{i\varphi}\psi_2(x)) = 0$$

and the interference is gone (at least on the average unless we obtain more knowledge about φ).

This example turns out to be generic. Systems behave quantum mechanically (that is show interferences, violate Bell-type inequalities, etc.) if their constituents have undisturbed ("coherent") relative phases while they start to behave classically as soon as these phase relations are lost.

As time evolution in quantum theory is unitary and thus invertible, of course nothing can get lost in a strict sense. The only thing that can happen is that this phase information is transferred to other degrees of freedom which

we will collectively call the "environment" (which specifically can include the macroscopic number of degrees of freedom of a measuring device).

This leaking of phase information is harder to prevent the more degrees of freedom the system can possible interact with and this in the end explains why we hardly see macroscopic interferences and the world around us appears to be classical: It is due to "decoherence" of the relative phases of Hilbert space vectors.

The individual contributions in the seminar investigate all these issues in much more detail.

1.3. What lies ahead?

The first two chapters on "algebraic quantum theory" and the the "Stone-von Neumann Theorem" deal with the question why the mathematical formalism of quantum physics is the way it is. We start without an a priori Hilbert space just with a set of the possible measuring apparatuses that we identify with the elements of an abstract (C^*)-algebra of observables A.

In this abstract setting, a state is nothing but a linear map $\omega : \mathcal{A} \to \mathbb{C}$ that maps observables to their expectation values. For the expectation value interpretation to be consistent it should be positive, that is, it should map positive operators of the form A^*A to non-negative numbers and it should be normalized $\omega(\mathbb{1}) = 1$. One might be worried that there is more information in a state about an operator than the expectation value like for example the variance. But of course the variance can itself be expressed as expectation values of the operator itself and its square.

As discussed in the contribution by Sebastian Seehars, from $\mathcal{A}$ and ω it is then possible to construct a Hilbert space on which the observables $\mathcal{A}$ are realized via a map $\pi(\mathcal{A})$, as (bounded) operators and which contains a vector $|\Omega\rangle$ such that the state ω is realized as expectation value $\omega(A) = \langle\Omega|\,\pi(A)\,|\Omega\rangle$. That is, the Hilbert space arises here as a representation of the algebra of observables. This representation is irreducible if the state ω is pure.

If the algebra is already an algebra of operators on a Hilbert space (which is in particular the case if it is a finite dimensional matrix algebra) then any

state can be written as $\omega(\mathcal{A}) = \mathrm{TR}(\rho A)$ with a density matrix (operator), that is a hermitian operator ρ that is positive and has $\mathrm{Tr}\,\rho = 1$.

Over and over in this seminar, we are dealing with the situation that the whole "world" we are describing falls into several parts like the system and the environment or the system and the measuring device or simply 'what we are interested in' and 'the rest'. Mathematically, this means we write the total Hilbert-space as a tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Then there are certain observables which only act on one part of the world, like only on the system or only on the measuring device. These are of the form $A = A_1 \otimes \mathbb{1}$ (or $B = \mathbb{1} \otimes B_2$ for that matter). Taken together, they form a sub-algebra $\mathcal{A}_1 \otimes \mathbb{1}$ and of course one can restrict any state ω to only this sub-algebra $\omega_1(A_1) = \omega(A_1 \otimes \mathbb{1})$. This state is then described by a reduced density matrix $\rho_1 = \mathrm{Tr}_{\mathcal{H}_2}(\rho)$ on $\mathcal{H}_1$ such that $\omega_1(A_1) = \mathrm{Tr}_{\mathcal{H}_2}(\rho_1 A_1)$.

Two remarks are in order. One should note that the partial trace does not arise because of some averaging over some non-observed environment but is a consequence of the choice only to test the state ω on observables of part of the world by restricting attention to the sub-algebra $\mathcal{A}_1$ Second, I would like to stress the the decomposition of $\mathcal{H}$ as a tensor product and the associated restriction of $\mathcal{A}$ to $\mathcal{A}_1$ s not set in the mathematical structure of $\mathcal{H}$ but is rather a choice of the observer who decides which observables to observe. In that sense, the reduced density matrix, which we will later find to have classical properties in many cases, is the result of a subjective choice of the restriction of attention rather than intrinsic in the world or its mathematical description.

Taking this abstract approach to observables and states clarifies an issue of locality that is often confused: The information about locality is encoded in the observables, not in the states. For the bipartite system above with $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ the local observables are those of the form $\mathcal{A}_1 \otimes \mathbb{1}$ or $\mathbb{1} \otimes A_2$ while the states or wave functions $\psi \in \mathcal{H}$ are always global objects. There is no such thing as "the state of particle 1" (where $\mathcal{H}_1$ is the Hilbert space of particle one), at least not when one makes measurements not of the form $\mathcal{A}_1 \otimes \mathbb{1}$, but that also measure properties of particle two. The only thing that exists is the global state of both particles that can contain correlations and entanglement. This is no particularity of quantum physics: Also classically, we are used to non-local correlations in the state: When in the morning

picking a random pair of socks from a drawer and putting them on, the color of the socks on both feet is correlated, for most people at least, even if the feet are at a space-like distance. Looking at the color of the left sock does not transfer any information about sock's color from the left to the right foot.

The contribution by Dennis Schimmel analyzes the construction of the Hilbert space and operators for the most common case that the algebra $\mathcal{A}$ is generated by variables x and p obeying canonical commutation relations (properly interpreted in terms of the Weyl algebra). It contains a proof of the Stone-von Neumann theorem stating that (up to unitary equivalence) there is only one irreducible Hilbert-space representation of these operators, the Schrödinger representation where x is the multiplication operator in position space and p is id/dx.

The next contribution by Mario Flory investigates the possibilities that arise when one considers a quantum mechanical system not only as a system by itself but one allows it to interact with further degrees of freedom that arise when one allows the Hilbert space to grow a further tensor factor describing this "environment". Specifically, in that case one can not only measure the eigenvalues of a hermitian operator acting on the system but one can generalize the projection valued measure of the spectral decomposition of the observable to positive operator valued measures (POVMs). As an example, it is explained how one can distinguish two not necessarily orthogonal Hilbert-space vectors (at the price of possibly reaching an inconclusive answer).

By allowing a system to interact with other degrees of freedom one can also significantly generalize the possible time evolutions: While for a closed system, only unitary transformations are possible, in connection with an environment any strictly positive, normalized linear map acting on the density matrix is possible (when considering this density matrix as a reduced density matrix of the world which undergoes itself a unitary time evolution). Again POVMs play a role.

1.3.1. Measurements

A measurement is the observation of properties of the state of a system by an observer. That is, the state of the observer after the measurement should

be influenced by the state of the system under observation. Otherwise, the measurement has not been effective. The "measurement problem" arises as we think of ourselves as observers as macroscopic, classical objects. Therefore, the quantum state of the system has to influence our classical state via the measurement. This appears to be problematic as the quantum state can be a superposition of two states of which the related classical states of the observer are macroscopically different. If the measurement is a linear process, a superposition on the system side should also lead to a superposition of the observer after the measurement, an operation that is not possible classically. The chapter on the measurement problem by Anupam Prasad Verdurmudi and David Jahn deals with this problem in detail.

1.3.2. lternatives to Copenhagen I: Decoherence

Modern approaches to the problem of the transition between the quantum and the classical realm as manifested in the measurement problem try to get rid of the collapse of the wave-function that is postulated by the traditional "Copenhagen"-interpretation for the moment of the measurement. The collapse turns a quantum state to a probabilistic classical state but at the price of an ad hoc assumption to give up the usual time evolution in favor of the projection to the eigenstate of the observed operator.

At the heart of the decoherence approach is the idea that the collapse mechanism is the effective description of a dynamical process that is described within the unitary evolution of quantum mechanics. The irreversibility of the collapse comes about since the collapse appears only when degrees of freedom attributed to "the environment" are not measured. The world consisting jointly of system and environment behaves quantum mechanically at all times but when the system alone is considered it looks as if a classical collapse has occurred.

Taking ψ_1 and ψ_2 to be basis vectors in a Hilbert space, the superposition

$$\frac{1}{\sqrt{2}}(\psi_1 + e^{i\varphi}\psi_2)$$

(with a possible relative phase φ as above) corresponds to the density matrix

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & e^{i\varphi} \\ e^{-i\varphi} & 1 \end{pmatrix}$$

Again, as above, averaging over φ removes the off-diagonal entries. The same happens upon enlarging the Hilbert-space by a tensor factor. If that has basis elements e_1 and e_2, then entangling the above state with this environment (for example a detector in a double slit experiment such that e_1 corresponds to the state of a detector that found the particle going through slit one while e_2 is the state of the detector that found the particle went trough slit two) leads to the state

$$\frac{1}{sqrt2}(\psi_1 \otimes e_1 + e^{i\varphi}\psi_2 \otimes e_2)$$

If one considers now the reduced density matrix of the system only by ignoring the environment (not reading out the detector) one finds

$$\rho' = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

In a sense, this mixed state is classical: It corresponds to adding the probabilities of finding ψ_1 or ψ_2 rather than the amplitudes as we found above. We could equivalently generate it by throwing a classical coin and then depending on the outcome take ψ_1 or ψ_2. The interferences, the signs of the quantumness of nature are gone as the relative phase φ has disappeared.[1]

By forgetting about the relative phase between ψ_1 and ψ_2, we have turned the quantum superposition into a classical probabilistic "either or" alternative. Still, half of the times one goes to check which is realized one finds the first but like in the weekly lottery nobody expects that any of the thirteen million other possible outcomes of drawing six balls from a collection of 49 influences the outcome that we perceive as realized. With respect to interference effects - the only difference between the quantum world and a probabilistic classical world - the state ρ' is on the classical side.

The resulting state ρ' is exactly the same that would have resulted from a

[1]This is somewhat curious: Strictly speaking, a pure state is not a non-zero vector ψ in the Hilbert-space but the equivalence class of all non-zero scalar multiples of ψ, the ray through ψ since to compute expectations one only needs the projector $\rho = |\psi\rangle\langle\psi|/|\psi|^2$ which is invariant under scalar multiplications. To form superpositions, however, only knowing the rays that interfere is not enough, one needs the actual vectors. With the rays one can only form classical statistical mixtures, for quantum superpositions one needs the vectors. The physical difference between the two only becomes important when taking into account other vectors/states to interfere with.

collapse of the wave function in the Copenhagen interpretation: With probability one half it is the state described by ψ_1 and with probability one half it is the state described by ψ_2.

The idea behind the "decoherence" approach to the quantum-classical transition is that this behavior is generic when a system is coupled to many unobserved degrees of freedom: The interaction with the environment turns off the off-diagonal entries in the density matrix thereby rendering a quantum state classical. This is explained in detail in the chapter on decoherence by Kostas Vavouranakis and Kosmas Kepesidis.

It should be stressed once more that this transition is subjective in the following sense: The diagonalization happens in the reduced density matrix that describes the state of the "system", while not taking into account the state of the environment. The world (defined as system plus environment) as a whole is a closed system and thus has a unitary time evolution under which pure quantum (superposition) states stay pure for all times. Only when restricting attention to the system it appears classical but the quantumness can be recovered by also observing the environment with which the system is entangled.

The decoherence chapter shows how the decoherence comes about dynamically by having the system interact with the environmental degrees of freedom and computes the time scale at which this happens. It also answers the question of the pointer basis: The density matrix being a hermitean operator can always be diagonalized. That is, one can always find a basis in which all off-diagonal entries are zero, we do not need decoherence to explain this. Thus, strictly speaking, the diagonalization is not the important fact about decoherence. It is the fact that in a particular given basis the matrix becomes diagonal. This basis is identified as the basis that the system-environment interaction measures: Take the the system-environment interaction in the interaction picture to be of the form

$$H = \sum_k H^{(k)} \otimes H_k$$

where k runs over the degrees of freedom of the environment and all $H^{(k)}$ act on the "system" while H_k only acts on the k^{th} degree of freedom of the environment. Decoherence happens now if all the $H^{(k)}$ commute and is

effective in the basis that simultaneously diagonalizes all these interaction Hamiltonians. For example, if the environment acts as spatial noise, that is if the interaction is realized by potentials $H^{(k)} = v(k)(x)$ decoherence happens in position space and the state at different x loses its relative phase exponentially fast.

1.3.3. Alternatives to Copenhagen II: Consistent Histories, Many Worlds and Bohmian Mechanics

The approach termed "Consistent Histories" is another interpretational framework originating in ideas about quantum cosmology. It aims to single out families of "histories" of events for which the classical addition of probabilities is possible while the time evolution is that of the Schrödinger equation. In the chapter by Hao Wu about this approach a mathematical criterion for consistency is developed.

The idea of the Many Worlds interpretation goes back to H. Everett. It can be closely linked to the decoherence approach discussed above and realizes the probabilistic final state ρ' after decoherence as an ensemble of worlds whose frequency realizes the probabilities on the diagonal of the reduced density matrix. The reality of these many parallel worlds can be debated. Whether they are just a figure of speech to give the probabilities a frequentist meaning or whether it makes sense to assume them to exists although being unobservable from the other worlds is a matter of opinion. This is explained in the chapter by Max Jeblick.

A radically different approach is taken by Bohmian mechanics. There, one insists on the positions of particles to be realistic, that is to have a value even at times when not observed. The lesson of the Bell inequalities is then that one has to sacrifice locality. This is done as follows: The wave function ψ is computed as always using the Schrödinger equation. For Hamiltonians of the form

$$H = \tfrac{1}{2}p^2 + V(x)$$

one can define a conserved probability current. The Bohmian idea is then to interpret this current as arising from the motion of particles via

$$\vec{j} = |\psi|^2 \vec{v}$$

that is to postulate particles with velocities

$$\dot{\vec{Q}} = \vec{v} = \mathrm{Im}\left(\frac{\nabla\psi}{\psi}\right)$$

An ensemble of particles that is distributed according to the density $\psi(t=0)|^2$ at initial time will be distributed like $|\psi(t)|^2$ at any time. The vector Q is interpreted as the position of the particles that are actually observed for example as black points on a photographic plate observing the spatial distributions of particles behind a double slit.

Locality is lost as the wave function depends on the positions of all particles and thus the equation giving the velocity of the ii^{th} particle depends on the simultaneous position of all other particles. All this is explained in the contribution by Henry Hanson and Franz Thoma.

The Bohmian interpretation is not followed by the majority of physicists. Points of criticism are the sacrifice of locality which many consider not natural and which makes it difficult to generalize this approach to the relativistic realm as well to the situation of quantum field theory where particles can be created and annihilated. Furthermore, it is not obvious what a certain observation measures, the wave function or the position Q. It is claimed that Q is observed when one sees the particle nature of the system, however, if one wants to describe the measurement quantum mechanically it should not measure v since the evolution of the wave function does not depend on Q (since there is no feedback of Q into the Schrödinger equation). In addition, for time-reversal invariant systems (effectively those without a magnetic field), one can take the Hamilton operator to be real and thus its eigenfunctions can be choses real as well (or with a constant phase). But for those $\dot{\vec{Q}} = 0$ and the particles donÕt move. This is the case even for the eigenfunctions of the harmonic oscillator and the hydrogen atom. In Bohmian quantum mechanics, the particles stand still even though there is a strong force acting on them. If the position Q is observable at all (otherwise its status as a physical ontological entity is doubtful) this should be experimentally checkable. Finally, the realism of the Bohmian approach only applies to position. Other observables like momentum or spin do not (and cannot) have realistic values unless measured. This is explained as "only positions are characteristic properties of particles" but is unclear what singles out these particular observables to receive this special treatment.

2. Algebraic quantum theory - Sebastian Seehars

In introductory lectures of quantum mechanics one usually starts by postulating that the state of a quantum system is given by an element of a (projective) Hilbert space called the 'wave function'. We can act on the system—and thus on its state—by a linear operator on this Hilbert space or by measuring an observable. In the algebraic approach to quantum theory, however, one starts with an abstract algebra of observables and then finds the Hilbert space as a derived concept, i.e. a representation of the observable algebra.

2.1. Algebras of Observables

In an n-dimensional HIlbert space, the algebra of observables consists of the complex $n \times n$ matrices, that is of $Mat(n \times n, \mathbb{C})$.

If the Hilbert space is infinite dimensional, it is conventional to use the algebra $\mathcal{B}(H)$ of bounded operators, where an operator A is called 'bounded' if its operator norm

$$\|A\| = \sup_{\psi \in \mathcal{H}, \psi \neq 0} \left(\frac{\|A\psi\|}{\|\psi\|} \right)$$

is finite. This boundedness condition is automatic in the finite dimensional case. In any case, $\mathcal{B}(H)$ has the structure of a C^*-algebra.

Definition 1 *A vector space A over $\mathbb{C}$ with operation $\circ : A \times A \to A$; $(a, b) \mapsto a \circ b$, such that the following holds:*

- $\circ$*-operation is bilinear:*

$$(\alpha a) \circ b = \alpha(a \circ b)$$
$$(a + a') \circ b = a \circ b + a' \circ b$$
$$a \circ (\alpha b) = \alpha(a \circ b)$$
$$(a \circ (b + b') = a \circ b + a \circ b'$$

- $\circ$*-operation is associative:* $a \circ (b \circ c) = (a \circ b) \circ c$

is called associative $\mathbb{C}$-algebra $\mathfrak{A}$.

Definition 2 *Let $\mathfrak{A}$ be an associative $\mathbb{C}$-algebra, $\|\cdot\|$ be a norm on the $\mathbb{C}$-vector space A. Let operation $* : A \to A;\ a \mapsto a^*$. Then $(A, \|\cdot\|, *)$ is called C^*-algebra $\mathfrak{A}$, if the following holds:*

- *$*$-operation is $\mathbb{C}$-antilinear :* $\lambda \in \mathbb{C}, a \in A : (\lambda a)^* = \bar{\lambda} a^*$,
- *$*$-operation is an involution:* $a^{**} = a$,
- $(ab)^* = b^* a^*$,
- *submultiplicity:* $\|ab\| \leq \|a\|\|b\|$,
- $\|a^*\| = \|a\|$,
- $\|a^* a\| = \|a^2\|$.

It is easy to see, that e.g., $(Mat(n \times n, \mathbb{C}), \dagger, \|\cdot\|_{op})$ is a C^*-algebra, where $\dagger$ is the hermitian conjugate and $\|\cdot\|_{op} := \sup_{\|\phi\|=1} \|A\phi\|$. In general, we can also replace $Mat(n \times n, \mathbb{C})$ by $\mathcal{B}(H)$ and still get a C^*-algebra.

2.2. States

In the algebraic framework, states are defined as expectation value functionals, mapping from the C^*-algebra to the complex numbers:

- *τ is normalized:* $\tau(\mathbb{1}) = 1$
- *τ is positive:* $\forall a \in \mathfrak{A} : \tau(a^* a) \geq 0$

One can show:

1. $\mathfrak{A} \times \mathfrak{A} \to \mathrm{C}; (a, b) \mapsto \tau(b^* a)$ is positive semi-definite, hermitian, sesquilinear form;
2. $|\tau(b^* a)|^2 \leq \tau(a^* a)\tau(B^* b)$;
3. $\tau(a^*) = \overline{\tau(a)}$;
4. $|\tau(a)|^2 \leq \tau(a^* a)$

Remark 4 *If we have an algebra $\mathfrak{A}$ of bounded operators on a Hilbert space, every density-operator ρ ($\rho \geq 0, Tr(\rho) = 1$) induces a state by $\tau : \mathfrak{A} \to \mathbb{C}; a \mapsto Tr(\rho a)$ (and vice versa in finite dimensions).*

2.3. Gelfand-Neumark-Segal Construction

In order to construct the connection of this abstract algebra notion to our standard quantum mechanics formalism, we still need a correspondence between the algebra and the linear operators. This connection will be given in terms of a "representation".

Definition 5 *Let $\mathfrak{A}, \mathfrak{B}$ be algebras with involution and unit $\mathbb{1}$. A map $\pi : \mathfrak{A} \to \mathfrak{B}$ is called $*$-homomorphism, if the following holds:*

- *π is $\mathbb{C}$-linear map: for $\lambda \in \mathbb{C} : \pi(\lambda a) = \lambda\pi(a)$*
- $\pi(ab) = \pi(a)\pi(b)$
- $\pi(a^*) = (\pi(a))^*$

Definition 6 *A representation is a $*$-homomorphism $\pi : \mathfrak{A} \to \mathcal{L}(\mathcal{D})$, where $\mathfrak{A}$ is C^*-algebra, $\mathcal{D}$ is dense in $\mathcal{H}$, $\mathcal{H}$ is Hilbert space, $\mathcal{L}(\mathcal{D})$ is the space of linear operators on $\mathcal{D}$ and $\pi(\mathbb{1}) = \mathbb{1}$.*

Construct scalar product and Hilbert space

From section 2.2 we know that the map $(a, b) \mapsto \tau(a^*b)$ is only positive semi-definite. But for a scalar product, we need positive definiteness. Therefore define the nullspace $\mathfrak{N}_\tau$:

$$\mathfrak{N}_\tau := \{a \in \mathfrak{A} | \tau(a^*a) = 0\}$$

and the quotient-set $\mathfrak{A}/\mathfrak{N}_\tau$, where this is a set of equivalence classes $[a]$:

$$[a] = \{b \in \mathfrak{A} | \exists n \in \mathfrak{N}_\tau : b = a + n\}$$

Let?s now pick two equivalence classes $[a]$ and $[b]$ and two elements of the nullspace n and m and evaluate the following quantity:

$$\tau((a+n)^*(b+m)) = \tau(a^*b) + \tau(a^*m) + \tau(n^*b) + \tau(n^*m)$$

For τ to be a well defined scalar product on the quotient set, we therefore need that the last three terms on the right hand side vanish. But by employing the Cauchy-Schwarz inequality $|\tau(a^*b)|^2 \leq \tau(a^*a)\tau(b^*b)$, we can easily see that:

$$|\tau(a^*m)|^2 \leq \tau(a^*a)\tau(m^*m) = 0$$

Similar equations also hold for the other two terms.

Now, we can define the scalar product $\langle\cdot,\cdot\rangle_\tau$:

$$\langle\cdot,\cdot\rangle_\tau : \mathfrak{A}/\mathfrak{N}_\tau \times \mathfrak{A}/\mathfrak{N}_\tau \to \mathbb{C}; ([a],[b]) \mapsto \tau(a^* b)$$

and the Hilbert space $\mathcal{H}_\tau$:

$$\mathcal{H}_\tau = (\mathfrak{A}/\mathfrak{N}_\tau, \langle\cdot,\cdot\rangle_\tau)$$

Construct representation

There is a general way to define the representation π_τ:

$$\pi_\tau : \mathfrak{A} \to \mathcal{L}(\mathcal{H}_\tau); \underbrace{\pi_\tau(a)}_{\in\mathcal{L}(\mathcal{H}_\tau)} \underbrace{[b]}_{\in\mathcal{H}_\tau} := \underbrace{[ab]}_{\in\mathcal{H}_\tau}$$

The question arises, whether or not the $[ab]$ notation is well defined, whether $[ab] = [a'b]$ if $[a] = [a']$. To see this, pick $b_1, b_2 \in \mathfrak{A}$, $n \in \mathfrak{N}_\tau$, such that $b_1 = b_2 + n$. Now take another $a \in \mathfrak{A}$:

$$ab_1 = ab_2 + an \Leftrightarrow ab_1 \sim ab_2, \text{ if } an \in \mathfrak{N}_\tau$$

Check if $an \in \mathfrak{N}_\tau$:

$$an \in \mathfrak{N}_\tau \Leftrightarrow \tau((an)^* an) \overset{!}{=} 0$$

$$|\tau((an)^* an)|^2 = |\tau(n^* a^* an)|^2 = |\tau((a^* an)^* n)|^2 \leq \tau((a^* an)^* a^* an) \underbrace{\tau(n^* n)}_{=0} = 0$$

Also, we define a unit vector $|\Omega\rangle = [\mathbb{1}]$. Then one can show:

$$\tau(a) = \tau(\mathbb{1}^* a \mathbb{1}) = \langle[\mathbb{1}], [a\mathbb{1}]\rangle = \langle\Omega_\tau, a\Omega_\tau\rangle$$

Definition 7 *Let τ be a state of the C^*-algebra $\mathfrak{A}$. The representation $(\mathcal{H}, \langle\cdot,\cdot\rangle, \pi_\tau, \Omega_\tau)$ is called* GNS-representation.

2.4. Example

Consider $\mathfrak{A} = Mat(2\times 2, \mathbb{C})$ with state $\tau : \mathfrak{A} \to \mathbb{C}; a \mapsto \mathrm{Tr}(\rho a)$, where ρ is a density matrix.

First look at $\mathfrak{N}_\tau$: for $a \in \mathfrak{N}_\tau$ the following holds:

$$\tau(a^*a) = 0 \Leftrightarrow \mathrm{Tr}(\rho a^* a) = 0$$

Since the trace is invariant under a change of basis, we can choose ρ to be diagonal:

$$\rho = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix}$$

$$\begin{aligned} \Rightarrow 0 &= \mathrm{Tr}\left(\begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix}\begin{pmatrix} a_1^* & a_3^* \\ a_2^* & a_4^* \end{pmatrix}\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}\right) \\ &= \rho_1(|a_1|^2 + |a_3|^2) + \rho_2(|a_2|^2 + |a_4|^2) \end{aligned}$$

(a) pure state

For a pure state, we can choose w.l.o.g.: $\rho_1 = 1, \rho_2 = 0 : \rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

$$\Rightarrow 0|a_1|^2 + |a_3|^2 \Leftrightarrow a_1 = a_3 = 0$$

$$\Rightarrow \mathfrak{N}_\tau = \{a \in \mathfrak{A} | a_1 = a_3 = 0\}$$

$$\mathfrak{A}/\mathfrak{N}_\tau \ni [a] : a \sim b = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} + \begin{pmatrix} 0 & n_2 \\ 0 & n_4 \end{pmatrix}$$

Out of every equivalence-class $[a]$, choose $\begin{pmatrix} a_1 & 0 \\ a_3 & 0 \end{pmatrix}$ to represent the class and think of it as a $\mathbb{C}^2$-vector.

The scalar-product $\langle \cdot, \cdot \rangle_\tau$ is then defined by:

$$\langle \cdot, \cdot \rangle_\tau : \mathfrak{A}/\mathfrak{N}_\tau \times \mathfrak{A}/\mathfrak{N}_\tau \to \mathbb{C}; ([a], [b]) \mapsto \tau(a^*b)$$

$$\tau(a^*b) = \mathrm{Tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a_1^* & a_3^* \\ 0 & 0 \end{pmatrix}\begin{pmatrix} b_1 & 0 \\ b_3 & 0 \end{pmatrix}\right) = a_1^* b_1 + a_3^* b_3 = \begin{pmatrix} a_1^* & a_3^* \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_3 \end{pmatrix}$$

It follows that the Hilbert space takes the form $\mathcal{H}_\tau = (\mathfrak{A}/\mathfrak{N}_\tau, \langle \cdot, \cdot \rangle_\tau)$. Think of the Hilbert space as $\mathbb{C}^2$ with its regular scalar-product.

The linear operators on $\mathcal{H}_\tau$ are now defined by the representation π_τ:

$$\pi_\tau : \mathfrak{A} \to \mathcal{L}(\mathcal{H}_\tau); \pi_\tau(a)[b] = [ab]$$

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}\begin{pmatrix} b_1 & 0 \\ b_3 & 0 \end{pmatrix} = \begin{pmatrix} a_1b_1 + a_2b_3 & 0 \\ a_3b_1 + a_4b_3 & 0 \end{pmatrix} = a\begin{pmatrix} b_1 \\ b_3 \end{pmatrix}$$

$$\Rightarrow \pi_\tau(a)[b] = a\begin{pmatrix} b_1 \\ b_3 \end{pmatrix}$$

and we can think of elements in $\mathcal{L}(\mathcal{H}_\tau)$ as two by two matrices.

(b) mixed state

Now use $0 < \rho_1, \rho_2 < 1$ as density matrix. Then the condition for $a \in \mathfrak{N}_\tau$ yields:

$$a_1 = a_2 = a_3 = a_4 = 0 : \mathfrak{N}_\tau = \{0\}$$

In this case, $\mathfrak{A}/\mathfrak{N}_\tau = \mathfrak{A}$ and the scalar product takes the form:

$$\langle \cdot, \cdot \rangle_\tau : \mathfrak{A} \times \mathfrak{A} \to \mathbb{C}; ([a], [b]) \mapsto \tau(a^* b)$$

$$\tau(a^* b) = \mathrm{Tr}(\rho a^* b) = \rho_1(a_1^* b_1 + a_3^* b_3) + \rho_2(a_2^* b_2 + a_4^* b_4)$$

and with A being the underlying vector space of the C^*-algebra $\mathfrak{A}$, the Hilbert space is $\mathcal{H}_\tau = (A, \langle \cdot, \cdot \rangle_\tau)$.

The representation π_τ acts as:

$$\pi_\tau : \mathfrak{A} \to \mathcal{L}(\mathcal{H}_\tau); \pi_\tau(a)[b] = [ab] = a \cdot b$$

and therefore is a regular matrix multiplication.

The idea is now to rewrite these equations to see how the "mixedness" of the state translates to a "mixedness" of the Hilbert space. Therefore, write the states $b \in \mathcal{H}_\tau$ as 4-vectors:

$$|b\rangle = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}$$

As a scalar product, one has:

$$\langle a, b \rangle = \begin{pmatrix} a_1^* & a_2^* & a_3^* & a_4^* \end{pmatrix}\begin{pmatrix} \rho_1 & 0 & 0 & 0 \\ 0 & \rho_1 & 0 & 0 \\ 0 & 0 & \rho_2 & 0 \\ 0 & 0 & 0 & \rho_2 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}$$

$$= \rho_1 \langle \begin{pmatrix} a_1 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_3 \end{pmatrix} \rangle + \rho_2 \langle \begin{pmatrix} a_2 \\ a_4 \end{pmatrix}, \begin{pmatrix} b_2 \\ b_4 \end{pmatrix} \rangle$$

and the operator $\pi_\tau(a)$ acting on the state $|b\rangle$ reads:

$$a\,|b\rangle = \left(\begin{array}{c|c} a & 0 \\ \hline 0 & a \end{array}\right)$$

where the matrix is a block-diagonal matrix. Now let's define the unit vector $|\Omega\rangle$:

$$|\Omega\rangle = [\mathbb{1}] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow \langle a\rangle = \langle \Omega, a\Omega\rangle = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \rho_1 & 0 & 0 & 0 \\ 0 & \rho_1 & 0 & 0 \\ 0 & 0 & \rho_2 & 0 \\ 0 & 0 & 0 & \rho_2 \end{pmatrix} \left(\begin{array}{c|c} a & 0 \\ \hline 0 & a \end{array}\right) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow \langle a\rangle = \rho_1 \begin{pmatrix} 1 & 0 \end{pmatrix} a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \rho_2 \begin{pmatrix} 0 & 1 \end{pmatrix} a \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

This can be interpreted as evaluating the expectation value of a in the pure states $\begin{pmatrix} 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \end{pmatrix}$ first and then calculating their weighted average.

2.5. Product States

Consider composed systems 1 and 2 with respective Hilbert spaces $(\mathcal{H}_1, \langle\cdot,\cdot\rangle_1)$ and $(\mathcal{H}_2, \langle\cdot,\cdot\rangle_2)$. The product space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ has scalar-product:

$$\langle x \otimes y, x' \otimes y'\rangle = \langle x, x'\rangle_1 \langle y, y'\rangle_2$$

and linear operators take the form:

$$a \in \mathcal{L}(\mathcal{H}_1), b \in \mathcal{L}(\mathcal{H}_2) : (a \otimes b)(|x \otimes y\rangle) = |ax\rangle \otimes |by\rangle$$

Now we want to carry this over to the algebraic formulation. Given two representations $(\mathcal{H}_1, \phi)$ and $(\mathcal{H}_2, \psi)$, where $\phi : \mathfrak{A} \to \mathcal{L}(\mathcal{H}_1)$ and $\psi : \mathfrak{B} \to \mathcal{L}(\mathcal{H}_2)$, then there exists the representation $\pi : \mathfrak{A}\otimes\mathfrak{B} \to \mathcal{L}(\mathcal{H}_1\otimes\mathcal{H}_2)$; $\pi(a\otimes b) = \phi(a) \otimes \psi b$. In a similar way, one can define a norm on $\mathfrak{C} = \mathfrak{A} \otimes \mathfrak{B}$. (This is of course just a hand-waving argument and has to be made mathematically rigorous, but I don't want to do this here.)

Let's consider states on such a C^*-tensor product $\mathfrak{C} = \mathfrak{A} \otimes \mathfrak{B}$. The most naive way to define a linear functional $\tau : \mathfrak{C} \to \mathbb{C}$ is the following: given $\mu : \mathfrak{A} \to \mathbb{C}$ and $\nu : \mathfrak{B} \to \mathbb{C}$ define the *product state* τ as:

$$\tau : \mathfrak{C} \to \mathbb{C}; \tau(a \otimes b) := \mu(a)\nu(b)$$

If we want to measure only on one part of the product space, say on $\mathfrak{A}$, we have to use as observables c:

$$\mathfrak{C} \ni c = \underbrace{a}_{\in \mathfrak{A}} \otimes \underbrace{id}_{\text{identity on } \mathfrak{B}}$$

and they lead to states

$$\tau^A(a) = \tau(a \otimes id)$$

For the specific example discussed in section 4, this amounts to

$$\mathfrak{A} = \mathfrak{B} = Mat(2 \times 2, \mathbb{C}) : \mathfrak{C} = Mat(2 \times 2, \mathbb{C}) \otimes Mat(2 \times 2, \mathbb{C})$$

$$\tau^A(a) = \tau(c) = \mathrm{Tr}(\rho c) = \mathrm{Tr}(\rho(a \otimes id)) = \mathrm{Tr}(\tilde{\rho} a); \text{ where } \tilde{\rho} = \mathrm{Tr}_B \rho$$

So from our decision to ignore subpart $\mathfrak{B}$, we got a similar description as the one for $\mathfrak{A}$ being no product space, but with ρ getting replaced by $\tilde{\rho}$, where the partial trace over the subsystem $\mathfrak{B}$ is performed on ρ. The fact that the Hilbert space $\mathcal{H}$ splits nicely into the desired $\mathcal{H}_1 \otimes \mathcal{H}_2$ is not canonical, but arises from our restriction to the subalgebra with elements $a \otimes id$.

Other states on $\mathfrak{C}$ are:

- τ is a *correlated state*, if $\exists\, a \in \mathfrak{A}, b \in \mathfrak{B}$, such that $\tau(a \otimes b) \neq \tau^A(a)\tau^B(b)$,
- τ is a *decomposable*, if it can be expressed by a convex combination of product states
- and *entangled*, if not.

2.6. Further reading

For the real deal on C^*-algebras check out the paper of C. Baer and C. Becker[2] There is a sript of Klaus Fredenhagen (in German, though) on "Algebraische Quantenfeldtheorie", where the first chapter is relevant[10].

3. The Stone-von Neumann uniqueness theorem - Dennis Schimmel

3.1. The Canonical Commutation Relation (CCR)

If we build up Quantum Mechanics from the CCR ($[Q, P] = i$), we have not said anything about wavefunctions. Indeed, the concept of wavefunctions in L^2 and Q as multiplication operator and P as differential is but one special representation of the CCR (the "Schrödinger-representation"). But there is in principle no special reason to pick this representation instead of any other. In particular, if we pick any other representation, with a different Hilbert space, is that even physically equivalent to the Schrödinger-representation? The Stone-von Neumann Uniqueness Theorem states exactly this equivalence.

There are some technical difficulties, as the CCR implies that at least one of Q or P is unbounded. This means that we cannot use the weak C^*-algebra structure.

But there is an elegant way around this problem: we look at the exponentiated version of the CCR: Take two one-parameter groups $U(\alpha) = \exp(i\alpha P)$, $V(\beta) = \exp(i\beta Q)$ with $U(\alpha)V(\beta) = \exp(i\alpha\beta)V(\beta)U(\alpha)$.

3.2. Uniqueness Theorem

If $\{U(a)|a \in \mathbb{R}\}$ and $\{V(a)|a \in \mathbb{R}\}$ are (weakly) Lebesgue-measurable families of unitary operators acting irreducibly on a separable Hilbert space $\mathcal{H}$, such that

$$U(a)U(b) = U(a+b)$$

$$V(a)V(b) = V(a+b)$$

$$U(a)V(b) = \exp(iab)V(b)U(a)$$

If $\{\tilde{U}(a)|a \in \mathbb{R}\}$ and $\{\tilde{V}(a)|a \in \mathbb{R}\}$ are (weakly) Lebesgue-measurable families of unitary operators acting irreducibly on a separable Hilbert space $\mathcal{H}'$, such that

$$\tilde{U}(a)\tilde{U}(b) = \tilde{U}(a+b)$$

$$\tilde{V}(a)\tilde{V}(b) = \tilde{V}(a+b)$$

$$\tilde{U}(a)\tilde{V}(b) = \exp(iab)\tilde{V}(b)\tilde{U}(a)$$

hen there exists a Hilbert space isomorphism $W : \mathcal{H} \to \mathcal{H}'$, which intertwines the action of U and V , that is:

$$WU(a)W^{-1} = \tilde{U}(a)$$

$$WV(a)W^{-1} = \tilde{V}(a)$$

3.3. Remarks

1. One of these representations can be taken to be the Schrödinger-representation with $\mathcal{H} = L^2(\mathbb{R})$, $(U(\alpha)\psi)(x) = \psi(x-\alpha)$ and $(V(\beta)\psi)(x) = \exp(-i\beta x)\psi(x)$.

2. The result can be generalized to a finite number of degrees of freedom (in the sense that n particles in $\mathbb{R}^m$ have a finite number of degrees of freedom), but not to an infinite number (QFT).

3. The operators Q, P can be recovered by "differentiating", i.e.,

$$P = \frac{d}{dt} U(t)|_{t=0}$$

3.4. Sketch of Proof

The idea is to show that every representation can be built out of some linear subspaces of the Hilbert space and unitary operators acting on these. In this construction, an irreducible representation is completely characterized by a single vector and a single family of operators. The isomorphism between representations is then the identification of the vectors and the families of operators. The complete proof by J. v. Neumann can be found in [23] (in German).

3.4.1. Step 1:

Introduce the *Weyl-System* $S(\alpha, \beta) := \exp(-\frac{1}{2}i\alpha\beta)U(\alpha)V(\beta)$ and show certain properties via straightforward calculations using the assumptions about U and V and their commutator:

1. $S(\alpha, \beta)S(\gamma, \delta) = \exp(\frac{i}{2}(\alpha\delta - \gamma\beta))S(\alpha + \gamma, \beta + \delta)$

2. $S(0,0) = 1$, this follows directly from the above via setting $\alpha = \beta = \gamma = \delta = 0$

3. $(S(\alpha,\beta))^{-1} = S(-\alpha,-\beta)$, from above via setting $\gamma = -\alpha, \delta = -\beta$

4. $S^*(\alpha,\beta) = S(-\alpha,-\beta)$

Further, if we have an isomorphism between S and $\tilde{S}$, we obviously have an isomorphism between U, V and $\tilde{U}, \tilde{V}$ (Simply set $\alpha = 0$ to obtain U, $\beta = 0$ to obtain V.

3.4.2. Step 2:

Introduce $A(a) := \int d\alpha\, d\beta\, a(\alpha,\beta)S(\alpha,\beta)$ for $a \in L^1(\mathbb{R}^2)$ and calculate properties of A (We denote a as $K(A)$ (although we do not yet know, if this exists)):

1. $\alpha' A(a) = A(\alpha' a)$

2. $A^*(a) = A(\overline{a(-\alpha,-\beta)})$

3. $A(a)S(u,v) = A(\exp(\frac{i}{2}(\alpha v - \beta u))a(\alpha - u, \beta - v))$

4. $S(u,v)A(a) = A(\exp(\frac{-i}{2}(\alpha v - \beta u))a(\alpha - u, \beta - v))$

5. $A(a) + A(b) = A(a(\alpha,\beta) + b(\alpha,\beta))$

6. $A(a) \cdot A(b) = A\left(\int d\xi\, d\eta\, \exp\frac{i}{2}(\alpha\eta - \beta\xi a(\alpha - \xi, \beta - \eta)b(\xi,\eta)\right)$

7. $A(a) = 0$ implies $a = 0$.

   Since this is rather difficult to prove, I will give a short sketch of the important steps:

   $A(a) = 0$ implies $0 = S(-u,-v)A(a)S(u,v)$. From the above properties one can calculate:

   $$S(-u,-v)A(a)S(u,v) = A(\exp(i(\alpha v - \beta u))a(\alpha,\beta))$$

   This implies

   $$0 = \int d\alpha\, d\beta\, \exp(i(\alpha v - \beta u))a(\alpha,\beta)\langle S(\alpha,\beta)g, g\rangle \quad \forall f, g \in \mathcal{H}$$

[In principle, knowing functional analysis, we are done: The expression above can be read as the Fourier-transform of $a(\alpha,\beta)\langle S(\alpha,\beta)g,g\rangle$, and we know that this yields the desired result.]

Thus for $P = \sum c_{k,l}\exp\left(i(k\alpha + l\beta)\right)$, where we sum only over finitely many terms:

$$0 = \int d\alpha\, d\beta\, P(\alpha,\beta)a(\alpha,\beta)\langle S(\alpha,\beta)f,g\rangle$$

Extend this via dominated convergence to other, more general P, in particular a step-function. Then we obtain: $a = 0$ almost everywhere, which is the best we can hope for.

3.4.3. Step 3:

Consider now a special choice of a. With this choice we can build up $\mathcal{H}$ by using one vector, applying the S's and then taking linear combinations. Thus the whole representation is "squeezed" into one vector:

1. Define $\tilde{A}$ by $\tilde{A} = A(\exp\left(-\frac{1}{4}\alpha^2 - \frac{1}{4}\beta^2\right))$. For this $\tilde{A}$ we have $\tilde{A}^* = \tilde{A}, \tilde{A} \neq 0$ (as $A(a) = 0$ implies $K(A) = 0$) and $\tilde{A}S(u,v)\tilde{A} = 2\pi\exp\left(-\frac{1}{4}u^2 - \frac{1}{4}v^2\right)\tilde{A}$, in particular $\tilde{A}^2 = 2\pi\tilde{A}$.

2. Consider the "eigenvectors" of A, that is solutions of $\tilde{A}f = 2\pi f$. As a is bounded, the set $M := \{f \in \mathcal{H} | Af = 2\pi f\}$ is a closed linear subspace. Thus each element f of M can be written as $f = \tilde{A}(\frac{1}{2\pi}f)$.

3. A direct calculation gives:

   $$\langle S(\alpha,\beta)f, S(\gamma,\delta)g\rangle = \exp\left(-\tfrac{1}{4}(\alpha-\gamma)^2 - \tfrac{1}{4}(\beta-\delta)^2 + \tfrac{i}{2}(\alpha\delta - \beta\gamma)\right)\langle f,g\rangle$$

   for $f,g \in M$. This implies that for an orthonormal basis $\{\phi_n\}$ of M:

   $$\langle S(\alpha,\beta)\phi_n, S(\gamma,\delta)\phi_m\rangle = \exp\left(-\tfrac{1}{4}(\alpha-\gamma)^2 - \tfrac{1}{4}(\beta-\delta)^2 + \tfrac{i}{2}(\alpha\delta - \beta\gamma)\right)\delta_{nm}$$

4. Now we want to see which part of our total Hilbert space cannot be reached by applying some S to a vector inside of M (it will turn out that $Lin(SM)$ is the total Hilbert space). For this define

   $$P_n := \{h \in \mathcal{H} | \exists \alpha \in \mathbb{C} : \alpha h = S(\phi_n)\}$$

Then $SP_n = P_n$ (as $S^2(\alpha,\beta) \propto S(\gamma,\delta)$) and $S^{-1}(\alpha,\beta) \propto S(\gamma,\delta)$. Thus from a simple calculation (using $\tilde{A}^* = \tilde{A}$): $\forall f \in (Lin(P_n))^\perp : \tilde{A}f = 0$ (Consider $\langle Ag, f\rangle$ for some $g \in \mathcal{H}$ and $f \in (Lin(P_n))^\perp$. With $\tilde{A}^2 = 2\pi\tilde{A}$ we have $Ag \in M$. Using $\tilde{A}^* = \tilde{A}$ we obtain $\tilde{A}f = 0\ \forall g \in \mathcal{H}$). This implies that $(Lin(P_n))^\perp = \{0\}$ (as the properties of $\tilde{A}$ [point 4.3.1] also apply when we restrict $\tilde{A}$ to $(Lin(P_n))^\perp$ which is a sub-Hilbert space of $\mathcal{H}$ and a subrepresentation of the S). Thus $M = \mathcal{H}$ (This is not surprising, since $Lin(S\xi)$ is an irreducible representation for some $\xi \in M$ and $\mathcal{H}$ is an irreducible representation which contains $Lin(S\xi)$.

3.4.4. Step 4:

Since the above construction is completely general, this can be done for both representations. But in the "projections" the desired isomorphism between S and $\tilde{S}$ is obvious, since then each element of the Hilbert space and each operator S is characterized only by coefficients (which can be equated) and numbers α, β corresponding to the action of $S(\alpha,\beta)$ on the vector giving the one-dimensional subspace.

In formulas we do the following: M and thus $\mathcal{H}$ can be segmented by the P_n. So if we can identify the P_n from $\mathcal{H}$ with those from $\mathcal{H}'$ in a sensible way (preserving the Hilbert space and representation structure), we have found the isomorphism we were looking for. The Hilbert space structure, in particular the scalar product is conserved for vectors inside $Lin(SM)$ (as its result only depends on on the scalar product of ϕ_1 with ϕ_j and the parameters of the S). We define $f_{\alpha,\beta} := S(\alpha,\beta)\phi_n$. But then the isomorphism is obvious: identify the $f_{\alpha,\beta}$ from $\mathcal{H}$ with the $f'_{\alpha,\beta}$ from $\mathcal{H}'$. This identification preserves everything we need. If we do this identification only for a basis of $\mathcal{H}$, it is obvious that the result is unique for every vector.

For irreducible representations we thus actually obtain isomorphy.

Note: For the Schrödinger-Representation we have

$$\phi_1 = \phi_1(q) = \pi^{-1/4} \exp\left(-\frac{1}{2}q^2\right)$$

$$f_{\alpha,\beta} = f_{\alpha,\beta}(q) = \pi^{-1/4} \exp\left(-\frac{1}{2}q^2+\right) - (\alpha + i\beta)q + \left(-\frac{\alpha^2}{2} + \frac{i\alpha\beta}{2}\right)\Big)$$

3.4.5. Generalization to more degrees of freedom

For a system of $k \in N$ degrees of freedom, we have as CCR's:

$$[P_m, P_n] = 0 = [Q_m, Q_n]$$

$$[Q_m, P_n] = i\delta_{nm}$$

We can introduce one U_n for each P_n and one V_m for each Q_m. Setting

$$\begin{aligned} &S(\alpha_1, \cdots, \alpha_k, \beta_1, \cdots, \beta_k) \\ &\quad := \exp\left(-\tfrac{1}{2} i(\alpha_1\beta_1 + \cdots + \alpha_k\beta_k)\right) U(\alpha_1)\cdots U(\alpha_k) B(\beta_1)\cdots V(\beta_k) \end{aligned}$$

$$\begin{aligned} A := \int & d\alpha_1 \cdots d\alpha_k d\beta_1 \cdots d\beta_k \\ & \exp\left(-\tfrac{1}{4}(\alpha_1^2 + \cdots + \alpha_k^2 + \beta_1^2 + \cdots + \beta_k^2)\right) S(\alpha_1, \cdots, \alpha_k, \beta_1, \cdots, \beta_k) \end{aligned}$$

we can apply the same reasoning as above. This reasoning cannot be done for $k = \infty$ (as shown by Haag's theorem).

4. POVMs and superoperators - Mario Flory

In this talk, it was described how positive operator valued measures (POVMs) can be used to generalize our understanding of what a measurement actually is. The connection between ordinary POVMs and the more special, but also more familiar projection valued measures (PVMs) was discussed, and the important Neumark's theorem was stated. Later, we discussed how POVMs (under the name of Kraus operators) are connected to the time evolution of some subsystem of a larger Hilbert space. This helped us to refine our understanding of the Schrödinger's cat paradox.

4.1. Introduction

The kind of measurement that we are used to in the mathematical description of quantum mechanics is represented by a hermitian operator which is applied to quantum states. Via the spectral decomposition, such an operator can be described as a sum of mutually commuting projection operators, weighted with the possible outcomes of the measurement. Two important questions

which arise in the study of such measurements will be answered in this talk, and the answer will in both cases be: positive operator valued measures, or a little bit shorter: POVMs!

The first question is: Can we generalize our understanding of measurements? In particular, if we perform an ordinary measurement on a given Hilbert space, how can we describe the action of the measurement on a subsystem without referring to the whole Hilbert space? The answer to this question will be given in section 4.3 by *Neumark's theorem* which will tell us that on the subsystem, the action of a projective measurement will always be described by a POVM (which we will define in section 4.2).

The second question is: If the Schrödinger equation tells us that the evolution of a closed quantum system is governed by a unitary transformation, how can we describe the evolution of a subsystem? We will see in the sections 4.4 and 4.5 that the answer to this is that the evolution of the subsystem is described by so called *Kraus operators* which by definition also form a POVM.

4.2. Positive operator valued measures (POVMs)

4.2.1. Definition

On a finite dimensional Hilbert space $\mathcal{H}$ (which will be sufficient for this talk), a POVM is a set of operators $\{E_m\}\,(m = 1, 2, ...)$ with the following properties [18]:

$$E_m^\dagger = E_m \text{ for all } m$$

$$\langle\psi| E_m |\psi\rangle \geq 0 \text{ for all } m \text{ and all } |\psi\rangle \in \mathcal{H}$$

We say the E_m are positive operators.

$$\sum_m E_m = 1 \quad \text{This is called a } \textit{partition of unity}$$

The underlying philosophy is the following: Imagine we have a given measurement-apparatus (which shall be represented by the POVM), with a set of possible measurement- results. Then we assign a *POVM element* E_m to every possible outcome, and we choose the E_m so that the probability of the corresponding outcome for a state $|\psi\rangle$ is given by the expectation value $\text{Prob}(m) = \langle\psi| E_m |\psi\rangle$. The requirements that we imposed upon the POVM elements ensure that all these probabilities are positive and add up to one

[18]. Note that the condition of hermiticity is redundant if the condition of positivity is given, but above it was nonetheless separately stated for the sake of clarity. Before looking at an application-example, we will clarify some more mathematical details. Afterwards, in section 4.3 we will discuss Neumark's theorem and hence get a better understanding of the meaning of POVMs. For a detailed description of how any possible POVM on the 2-dimensional Hilbert space of photon-polarization can be implemented in an actual measurement, see [1]. For the definition of POVMs in the case of infinite-dimensional Hilbert spaces see [8].

4.2.2. Measurement operators

There always exists a set of *measurement operators* $\{M_m\}$ which describes the POVM $\{E_m\}$ via

$$E_m = M_m^\dagger M_m$$

We write $M_m = \sqrt{E_m}$. Such a set need not be unique, but for every POVM it is possible to find one, due to the positivity of the operators E_m. [18]

As we know, the probabilities for the different possible outcomes of the measurement are already given by the POVM elements, the measurement operators are only needed if we are interested in the new state after the measurement, which in the case of the result corresponding to E_m is given by ([18])

$$|\psi_m\rangle = \frac{M_m |\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}}$$

So obviously the state after the measurement depends on the result of the measurement, this corresponds to the *collapse of the wave function* in the Copenhagen interpretation of quantum mechanics. Please note that in the case of POVMs the word "collapse" may not be fully adequate, as the repeated application of a POVM-measurement on a certain quantum system will not always yield the same result again, in contrast to the projection valued measures discussed in the next section. Nevertheless, such a kind of (time) evolution of a quantum system is not unitary and not time-invertible, in contrast to the time evolution described by the Schrödinger equation.

In general, the initial state may be described by a density matrix. If we now imagine that a measurement has been performed, but we have forgotten the

result, then the state after the measurement has again to be described by a density matrix ([21])

$$\rho' = \sum_m M_m \rho M_m^\dagger$$

We will recover exactly the same equation later as (4.1) in a completely different context. Note that:

$$\mathrm{Tr}[\rho'] = \mathrm{Tr}\left[\sum_m M_m \rho M_m^\dagger\right] = \sum_m \mathrm{Tr}[M_m^\dagger M_m \rho] = \mathrm{Tr}\left[\sum_m E_m \rho\right] = \mathrm{Tr}[\rho]$$

4.2.3. Projection valued measures (PVMs)

Usually, an observable is described by a hermitian operator A with decomposition

$$A = \sum_m a_m P_m$$

in which a_m is the m^{th} eigenvalue, and P_m is the projection operator on the corresponding eigenspace of A, with the properties $P_m^\dagger = P_m, P_m P_n = \delta mn P_m$ and $\sum_m P_m = 1$. Such a PVM is a special case of a POVM with $M_m = E_m = P_m$. Note the following differences between POVMs and PVMs: The operators P_m are mutually orthogonal, while this need not be the case for the E_m. Therefore, the number of projection operators P_m in the decomposition of A is limited by the dimension of $\mathcal{H}$, while the number of E_m is unlimited. Also, if a PVM is applied to a certain quantum state repeatedly, it will always yield the same result, while this need not be true for a general POVM.[18]

4.2.4. Application example

Suppose Alice sends Bob Q-bits in one of the two states

$$|\psi_1\rangle = |0\rangle \quad \text{and} \quad |\psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

with the same probability of $1/2$. We know that Bob is principally not able to distinguish between $|\psi_1\rangle$ and $|\psi_2\rangle$ perfectly, as they are not orthogonal.

But Bob can use a POVM of the form

$$E_1 = \frac{\sqrt{2}}{1+\sqrt{2}} |1\rangle\langle 1|$$
$$E_2 = \frac{\sqrt{2}}{2(1+\sqrt{2})} (|0\rangle - |1\rangle)(\langle 0| - \langle 1|)$$
$$E_3 = 1 - E_1 - E_2$$

So now, if Bob gets the result corresponding to E_1 (E_2) he can be sure that he got the $|\psi_2\rangle$ ($|\psi_1\rangle$) from Alice because $E_1 |\psi_1\rangle = E_2 |\psi_2\rangle = 0$. If he gets the result E_3,the measurement has to be discarded, but in contrast to any possible PVM (like for example $\{|0\rangle\langle 0|, |1\rangle\langle 1|\}$, Bob exactly knows when he has to ignore the outcome of a measurement, and therefore never makes a mistake in distinguishing between the two possible states. [18]

How can this example be related to the usual projection valued measures that we are familiar with? It turns out that, by adding one more dimension to the Hilbert space, for example by coupling the measured state to a known *ancilla* state, our given POVM can be described as a PVM in the larger dimensional Hilbert space [21]. This statement can be generalized and is related to Neumark's theorem which will be dealt with in the next section.

In this particular example, we first need to calculate the spectral decompositions of the POVM elements [21]. It turns out that all of them have one eigenvalue = 0 and can therefore be written in the form $E_i = |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|$ with non normalized states $|\tilde{\psi}_i\rangle$. For $i = 1, 2$ this is just the form in which they are given, in the case $i = 3$ the fact that one eigenvalue is zero is noteworthy. The prefactors in the definitions of E_1 and E_2 have to be chosen as large as possible, in order to minimize the expectation value of E_3, and therefore also the probability to get an inconclusive result. But also, we have to ensure that E_3 is a positive operator, in order to be a POVM element. As it turns out, one eigenvalue of E_3 is zero because the prefactors of E_1 and E_2 are chosen optimally. If they were larger, the second eigenvalue of E_3 would be negative and the POVM-conditions would be violated.

Using the basis $|1\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|0\rangle = \begin{pmatrix}0\\1\end{pmatrix}$, we find

$$|\tilde{\psi}_1\rangle = \sqrt{\frac{\sqrt{2}}{1+\sqrt{2}}}\begin{pmatrix}1\\0\end{pmatrix}$$
$$|\tilde{\psi}_2\rangle = \sqrt{\frac{\sqrt{2}}{2(1+\sqrt{2})}}\begin{pmatrix}-1\\1\end{pmatrix}$$
$$|\tilde{\psi}_3\rangle = \sqrt{\frac{1}{\sqrt{2}}}\begin{pmatrix}\sqrt{2}-1\\1\end{pmatrix}$$

If we add one dimension to our Hilbert space, then we have to extend E_1, E_2 and E_3 to a set of 3-dimensional projection operators P_1, P_2 and P_3 which project onto mutually orthogonal subspaces of the Hilbert space, i.e., which mutually commute and therefore can be composed to an ordinary observable $A = \sum_m a_m P_m$. With the definition $P_i = |\phi_i\rangle\langle\phi_i|$, we now have to search for an orthonormal set of 3-dimensional vectors

$$\phi_i = \begin{pmatrix}\tilde{\psi}_i\\?_i\end{pmatrix}$$

with a yet to be determined third component (we abandon the bra-ket notation for a while). We know that if the rows of a square matrix form an orthonormal basis (ONB), then the columns do so as well. So we can take the rows of the 3×3-matrix (ϕ_1, ϕ_2, ϕ_3) to be a set of two known vectors v_1 and v_2 and a completely unknown vector $v_3 = (?_1, ?_2, ?_3)^T$. Indeed, the two vectors v_1 and v_2 are already orthonormal with respect to each other. This may seem surprising, but indeed it is a result of the requirement that our POVM has to be a partition of unity [21]. The construction of v_3 is now quite easy: we just take an arbitrary vector which is not in span (v_1, v_2) (like $(1,0,0)^T$) and then apply the Gram-Schmidt orthonormalization-procedure. In our case, this yields

$$v_3 = ((1+\sqrt{2})^{-1/2}, (1+\sqrt{2})^{-1/2}, 1 - sqrt2)^T$$

and therefore we find:

$$|\phi_1\rangle = \begin{pmatrix}\sqrt{\frac{\sqrt{2}}{1+\sqrt{2}}}\\0\\\frac{1}{\sqrt{2+\sqrt{2}}}\end{pmatrix} \quad |\phi_2\rangle = \begin{pmatrix}\frac{-1}{\sqrt{1+\sqrt{2}}}\\\frac{1}{\sqrt{2+\sqrt{2}}}\\\frac{1}{\sqrt{1+\sqrt{2}}}\end{pmatrix} \quad |\phi_3\rangle = \begin{pmatrix}\frac{\sqrt{2}-1}{\sqrt[4]{2}}\\\frac{1}{\sqrt[4]{2}}\\1-\sqrt{2}\end{pmatrix}$$

This finally yields

$$P_1 = \begin{pmatrix} \frac{\sqrt{2}}{1+\sqrt{2}} & 0 & \frac{\sqrt[4]{2}}{1+\sqrt{2}} \\ 0 & 0 & 0 \\ \frac{\sqrt[4]{2}}{1+\sqrt{2}} & 0 & \frac{1}{1+\sqrt{2}} \end{pmatrix}$$

$$P_2 = \begin{pmatrix} \frac{1}{2+\sqrt{2}} & \frac{-1}{2+\sqrt{2}} & \frac{-1}{\sqrt[4]{2}(1+\sqrt{2})} \\ \frac{-1}{2+\sqrt{2}} & \frac{1}{2+\sqrt{2}} & \frac{1}{\sqrt[4]{2}(1+\sqrt{2})} \\ \frac{-1}{\sqrt[4]{2}(1+\sqrt{2})} & \frac{1}{\sqrt[4]{2}(1+\sqrt{2})} & \frac{1}{1+\sqrt{2}} \end{pmatrix}$$

$$P_3 = \begin{pmatrix} \frac{3-2\sqrt{2}}{\sqrt{2}} & \frac{\sqrt{2}-1}{\sqrt{2}} & \frac{2\sqrt{2}-3}{\sqrt[4]{2}} \\ \frac{\sqrt{2}-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1-\sqrt{2}}{\sqrt[4]{2}} \\ \frac{2\sqrt{2}-3}{\sqrt[4]{2}} & \frac{1-\sqrt{2}}{\sqrt[4]{2}} & 3-2\sqrt{2} \end{pmatrix}$$

and the 2×2 parts in the upper left corners are exactly the representations of the E_i. The measurement of one of the states $|\psi_1\rangle = (0,1)^T$ or $|\psi_2\rangle = \frac{1}{\sqrt{2}}(1,1)^T$ with the given POVM in the original Hilbert space turns now out to be analogous to the measurement of the states $(0,1,0)^T$ or $\frac{1}{\sqrt{2}}(1,1,0)^T$ with the PVM given above in the extended 3-dimensional Hilbert space.

4.3. Neumark's theorem

Theorem: Any POVM on a Hilbert space $\mathcal{H}_A$ can be realized by a PVM in a larger Hilbert space $\mathcal{H}$ containing $\mathcal{H}_A$.

Please note: As mentioned in the introduction, the reverse is also true: Every PVM on a Hilbert space $\mathcal{H}$ leads to a POVM on any subspace $\mathcal{H}_A$. [18][21]

Therefore, if we want to understand the effect of measurements on a subsystem without referring to the whole Hilbert space, POVMs appear to be a necessary and indeed very useful generalization of PVMs.

4.3.1. Consider the case $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$

Suppose that we have given an initial state $\rho_{AB} = \rho_A \otimes \rho_B$ and a PVM described by the set of operators P_m acting on $\mathcal{H}$. It is our aim to describe the probability to get the result corresponding to index m only by making

use of operators and states on $\mathcal{H}_A$. We therefore define a set of operators F_m on $\mathcal{H}_A$ such that [21]:

$$\text{Prob}(m) = \text{Tr}_{AB}[P_m \rho_{AB}] = \text{Tr}_A[\text{Tr}_B[P_m \rho_{AB}]] \equiv \text{Tr}_A[F_m \rho_A]$$

It is easy to find such F_m by writing the above equation in terms of components:

$$\begin{aligned} \sum_{i,j} (F_m)_{ji} (\rho_A)_{ij} &= \sum_{i,j,\mu,\nu} (P_m)_{ji,\nu\mu} (\rho_B)_{\mu\nu} (\rho_A)_{ij} \\ \Rightarrow (F_m)_{ji} &= \sum_{\mu,\nu} (P_m)_{ji,\nu\mu} (\rho_B)_{\mu\nu} \end{aligned}$$

Here the indices $i, j(\mu, \nu)$ denote components corresponding to base vectors in $\mathcal{H}_A$ ($\mathcal{H}_B$). Based on this definition of the F_m, it is easy to compute the following properties:

- $F_m = F_m^\dagger$
- F_m is a positive operator aor all m
- $\sum_m F_m = \mathbb{1}_A$

Proof 1: $(F_m)^*_{ij} = \sum_{\mu,\nu} (P_m)^*_{ij,\mu\nu} (\rho_B)^*_{\mu\nu} = \sum_{\mu,\nu} (P_m)_{ij,\mu\nu} (\rho_B)_{\mu\nu} = (F_m)_{ji}$ □

Proof 2: We set the basis such that $\rho_B = \sum_\mu p_\mu |\mu_B\rangle \langle\mu_B|$. Then: $\langle\psi_A| F_m |\psi_A\rangle = \sum_\mu p_\mu (\langle\psi_A| \otimes \langle\mu_B|) P_m (|\psi_A\rangle \otimes |\mu_B\rangle) \geq 0 \ \forall \psi_A \in \mathcal{H}_A$ □

Proof 3: $(\sum_m F_m)_{ij} = \sum_{\mu,\nu} [\sum_m (P_m)_{ij,\mu\nu}] (\rho_B)_{\mu\nu} = \sum_{\mu,\nu} (\mathbb{1}_{AB})_{ij,\mu\nu} (\rho_B)_{\mu\nu} = (\text{Tr}[\mathbb{1}_{AB} \rho_B])_{ij} = (\mathbb{1}_A)_{ij}$ □

Therefore, the set F_m indeed forms a POVM. In the proofs we used the hermiticity of the P_m and ρ_B as well as the positivity of ρ_B and the fact that $\text{Tr}_B[\rho_B] = 1$. [21]

4.4. Superoperators and unitary evolution

Postulate: The time evolution of a closed quantum system is governed by unitary transformations. [18]

But this need not be true in a subsystem A of a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$! Suppose our initial state is given by a density matrix

$$\rho_{AB} = \rho_A \otimes |0\rangle \langle 0|_B$$

After the unitary evolution defined by a unitary operator U_{AB} on $\mathcal{H}$, the new state in subsystem A is given by

$$\rho'_A = \mathrm{Tr}_B \left(U_{AB} \rho_{AB} U_{AB}^\dagger \right) = \sum_\mu \underbrace{\langle \mu_B | U_{AB} | 0 \rangle}_{\equiv M_\mu} \rho_A \underbrace{\langle 0 | U_{AB}^\dagger | \mu_B \rangle}_{\equiv M_\mu^\dagger}$$

where the vectors $|\mu_B\rangle$ form an orthonormal basis in $\mathcal{H}_B$. Now we define a superoperator $\$(\rho_A) \equiv \rho'_A$ describing the time evolution of ρ_A entirely in terms of operators acting on $\mathcal{H}_A$ by

$$\$(\rho_A) = \sum_\mu M_\mu \rho_A M_\mu^\dagger \tag{4.1}$$

in this example. The M_μ are called *Kraus operators* and satisfy

$$\left(M_\mu^\dagger M_\mu \right) \text{ is hermitian and positive for all } \mu$$

$$\sum_\mu M_\mu^\dagger M_\mu = \mathbb{1}_A$$

therefore they are the measurement operators of a POVM [21]. This is how unitary time evolution acts on the subsystem A.

Now, we want to investigate the opposite question: What are the allowed time- evolutions $\rho_A \mapsto \rho'_A = \$(\rho_A)$ of systems (that are possibly only subsystems of the whole universe) if the criterium is that physical states are mapped to physical states, i.e., that density matrices (positive and normalized) are mapped again to density matrices (rather than general matrices). In particular, we should at least demand[21, 5]

(0) $\$(\rho_A)$ is linear

(1) ρ_A is hermitian $\Rightarrow \$(\rho_A)$ is hermitian

(2) $\mathrm{Tr}_A(\rho_A) = 1 \Rightarrow \mathrm{Tr}(\$(\rho_A)) = 1$

(3a) ρ_A is positive $\Rightarrow \$(\rho_A)$ is positve

It turns out that condition (3a) is not sufficient: It does not imply that if we trivially extend the Hilbert space by another tensor factor and leave the state there untouched, positivity is preserved on the whole state. Thus, we have to demand it explicitly (this condition is called *complete positivity*):

(3b) ρ_A and ρ_C are positive $\Rightarrow (\$ \otimes \mathbb{1})(\rho_A \otimes \rho_C) = \$(\rho_A) \otimes \rho_C$ is positive for any extension $\mathcal{H}_C$

There are in fact examples of operators which satisfy (3a) but not (3b), as for example $\$(\rho) = \rho^T$ [21]. Also, whether (0) is a necessary condition or not is a much more subtle question than it may have become clear up to now, for a detailed discussion see [21].

Kraus representation theorem: Every operator $\$$ satisfying (0), (1), (2) and (3b) (called a *super operator*) has an operator-sum representation (4.1) and therefore also a unitary representation on a larger Hilbert space $\mathcal{H}$ [21][5]. Thus any time evolution that maps density matrices to density matrices can be thought of as arising as the evolution of a subsystem of a possibly larger Hilbert space that evolves unitarily.

Proof: In order to prove this important result, we first introduce the so called *relative state method* [21]: We suppose to have given a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim(\mathcal{H}_B) \geq \dim(\mathcal{H}_A)$ and a state $|\psi\rangle_{AB} = \sum_i |i_A\rangle \otimes |i'_B\rangle \in \mathcal{H}$, where the vectors $|i_A\rangle$ and $|i'_B\rangle$ are the first $\dim(\mathcal{H}_A)$ base-vectors in the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. In order to express $|\psi\rangle_{AB}$ in this simple way (with a single sum $\sum_i$ instead of a double sum) we have to choose the orthonormal basis $|i'_B\rangle$ of $\mathcal{H}_B$ appropriately, the possibility ?to do this is ensured by the *Schmidt decomposition theorem.*

Now if $|\phi_A\rangle = \sum_i a_i |i_A\rangle \, in \mathcal{H}_A$, we define the *relative state* $|\phi^*_B\rangle = \sum_i a^*_i |i'_B\rangle \in \mathcal{H}_B$ such that $|\phi_A\rangle = \langle \phi^*_B | \, \psi_{AB}\rangle$. This is called a partial scalar product. So for

$$(O_A \otimes \mathbb{1}_B) |\psi_{AB}\rangle = \sum_i O_A |i_A\rangle \otimes |i'_B\rangle$$

we find by applying the relative state method:

$$O_A |\phi_A\rangle = \langle \phi^*_B | \, (O_A \otimes \mathbb{1}_b) |\psi_{AB}\rangle$$

Now we apply this method in the proof of the Kraus-representation-theorem. Our superoperator \$ acts in the following way:

$$(\$_A \otimes \mathbb{1}_B) \underbrace{\rho_{AB}}_{\equiv|\psi_{AB}\rangle\langle\psi_{AB}|} = \underbrace{\sum_\mu q_\mu |\Phi_{\mu AB}\rangle\langle\Phi_{\mu AB}|}_{=\rho'_{AB}}$$

Note that the pure state form $\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$ can, for a given density matrix in subspace $\mathcal{H}_A$, always be achieved by choosing the dimension of $\mathcal{H}_B$ large enough. Via the relative state method, we obtain

$$\$_A(|\phi_A\rangle\langle\phi_A|) = \langle\phi_b^*|(\$_A \otimes \mathbb{1}_B)\rho_{AB}|\phi_B^*\rangle = \sum_m u \underbrace{\sqrt{q_\mu}\langle\phi_B^*|\Phi_{\mu AB}\rangle}_{\equiv M_\mu|\phi_A\rangle} \underbrace{\langle\Phi_{\mu AB}|\phi_B^*\rangle\sqrt{q_\mu}}_{\langle\phi_A|M_\mu^\dagger}$$

These M_μ have the following properties:

- they define a linear map $|\phi_A\rangle \to \sqrt{q_\mu}\langle\phi_B^*|\Phi_{\mu AB}\rangle$
- $\$_A(|\phi_A\rangle\langle\phi_A|) = \sum_\mu M_\mu|\phi_A\rangle\langle\phi_A|M_\mu^\dagger \;\forall\, |\phi_A\rangle \in \mathcal{H} \Rightarrow \$_A(\rho_A) = \sum_\mu M_\mu\rho_A M_\mu^\dagger$ if \$ is linear,
- $\sum_\mu M_\mu M_\mu^\dagger = \mathbb{1}_A$, because $\$_A$ is trace preserving for any ρ_A. $\Rightarrow \$_A$ has an operator-sum - respectively Kraus-representation (4.1)[21].

Now we still have to prove that from a operator-sum representation there follows a unitary representation on a larger Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_C$ such that $M_\mu = \langle\mu_C|U_{AC}|0_C\rangle$. Note that the Hilbert space $\mathcal{H}_C$ with which we extend $\mathcal{H}_A$ need not be equal to the Hilbert space $\mathcal{H}_B$ which we used in the proof of the Kraus-representation. We define a operator U_{AC} such that

$$U_{AC}(|\phi_A\rangle \otimes |0_C\rangle) \equiv \sum_\mu M_\mu|\phi_A\rangle \otimes |\mu_C\rangle$$

with some normalized state $|0_C\rangle \in \mathcal{H}_C$. It can easily be checked that the operator U_{AC} that is defined via this equation preserves the inner product:

$$(\langle 0_C| \otimes |\phi_{2A}\rangle U_{AC}^\dagger)(U_{AC}|\phi_{1A}\rangle \otimes |0_C\rangle) = ... = \langle\phi_{2A}|\phi_{1A}\rangle$$

From this it follows that U_{AC} can be extended to be a unitary operator on the whole Hilbert space, and indeed from the so defined operator we can obtain back the Kraus operators M_μ by performing the same steps as above.

[21] □

So, in this section we found that while the evolution of a closed quantum system is governed by unitary transformations (which can always be inverted, such that the evolution is reversed), the evolution of a subsystem is governed by much more general superoperators which are closely related to the formalism of POVMs. One important fact that has not yet been explicitly stated is that in general the inverse of a superoperator $ need not exist, or be again a superoperator itself. Therefore, it may not be possible to reverse the time evolution in a subsystem of a larger quantum system, so to say the subsystem "forgets" its past. This resolves some philosophical problems of quantum mechanics, as a unitary time evolution can be reversed, but the often postulated collapse of the wavefunction is an irreversible process. [21]

Also, it should be pointed out that the description of the time evolution of a sub- system discussed above leads to a large variety of possible transformations that can be expressed via superoperators, unitary as well as non-unitary ones. But of course there cannot be superoperators describing a time evolution which violates a basic quantum mechanical principle like for example the *no-cloning theorem.*

So far, we have only used superoperators mapping density matrices of a (sub)space $\mathcal{H}_A$ on density matrices of the same Hilbert space $\mathcal{H}_A$. But this is not the most general case possible. If we have a given Hilbert space $\mathcal{H}$ with two distinct subspaces $\mathcal{H}_A$ and $\mathcal{H}_{A'}$, then it is possible (if a unitary transformation on $\mathcal{H}$ is given) to find a corresponding superoperator $ which maps a density matrix ρ_A (defined on $\mathcal{H}_A$) on a density matrix ρ'_A (defined on $\mathcal{H}_{A'}$). Imagine for example a large laboratory (mathematically described by $\mathcal{H}$) in which the state of a single Q-bit ($\mathcal{H}_A$) at the beginning of an experiment determines the state of some apparatus ($\mathcal{H}_{A'}$) after the experiment. Then there exists a superoperator $ mapping the initial state in $\mathcal{H}_A$ on the final state in $\mathcal{H}_{A'}$. In the following section, when we will have a look on Schrödinger's cat, we will consider the cat together with the nucleus as one subsystem, but as shown above we could as well consider them to be two distinct subsystems.

4.5. Schrödinger's cat?s revival

I assume the reader to be familiar with the basic concept of the paradox known as "Schrödinger's cat". The cat together with the nucleus can be described roughly by a two dimensional Hilbert space with base-vectors $|\uparrow\rangle$ which means the nucleus is not decayed, the cat is still alive, and the vector $|\downarrow\rangle$ which corresponds to a decayed atom and a dead cat. Now we assume this subsystem to be in contact with the environment which acts as some kind of thermal bath. We now have to come up with a reasonable model of the (unitary) evolution of the whole system. For this purpose, we will describe the environment via a 3-dimensional Hilbert space with basevectors $|0\rangle_E$, $|1\rangle_E$, $|2\rangle_E$. With a high probability (for sufficiently small timesteps), no molecule is scattered off the cat, so that the whole system stays in the same state, but with a small probability p a molecule is scattered off the cat and the environment switches to another state, which depends on the state of the cat. Also, the state of the cat will never be changed by interaction with the environment. For example, if cats could be killed by the hit of an air-molecule, this would be a severe shortcoming in Darwinian evolution. In formulas, this model reads:

$$|\uparrow\rangle|0\rangle_E \to \sqrt{1-p}\,|\uparrow\rangle|0\rangle_E + \sqrt{p}\,|\uparrow\rangle|1\rangle_E$$
$$|\downarrow\rangle|0\rangle_E \to \sqrt{1-p}\,|\downarrow\rangle|0\rangle_E + \sqrt{p}\,|\downarrow\rangle|2\rangle_E$$

As can easily be checked, the evolution described by this rules is indeed inner product preserving and can therefore be extended to a unitary transformation. Via the usual steps, we find the Kraus operators on the 2-dimensional Hilbert space of the cat

$$M_0 = \sqrt{1-p}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \quad M_1 = \sqrt{p}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} \quad M_2 = \sqrt{p}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$$

They are indeed the measurement operators of a POVM, as again can easily be checked. Suppose the initial (pure!) state of our feline subsystem is

$$\rho_A = \frac{1}{2}(|\uparrow\rangle - |\downarrow\rangle)(\langle\uparrow| - \langle\downarrow|) = \frac{1}{2}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}$$

Now via the Kraus-representation, we find that after one timestep the new density matrix is

$$\$(\rho_A) = \ldots = \frac{1}{2}\begin{pmatrix}1 & -(1-p)\\ -(1-p) & 1\end{pmatrix}$$

This looks like a promising beginning! Now, we define a scattering-rate Γ and a timestep $\Delta t \ll 1$ such that $p = \Gamma \Delta t$. Then after a finite time $t = n\Delta t$ with large n, the cat has acquired the state

$$\rho'_A = \$^n(\rho_A) = \frac{1}{2}\begin{pmatrix} 1 & -\left(1-\frac{\Gamma t}{n}\right)^n \\ -\left(1-\frac{\Gamma t}{n}\right)^n & 1 \end{pmatrix} \to \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

So obviously, via interaction with the environment, the initial *pure state* (which describes the seemingly paradoxical quantum-superposition of life and death of a macroscopic object) has become a *mixed state*! Of course, we still don't know whether the cat will be alive or dead before we have opened the box, but the kind of uncertainty expressed by a mixed state density matrix is classical rather than the kind of uncertainty expressed by a pure state density matrix. [21]

5. The measurement problem - David Jahn and Anupam Prasad

5.1. Introduction

The purpose of a measurement is, in general, to determine properties of the physical system under investigation. In Quantum Mechanics this simple statement should be overthought. What are "properties" of a system? In classical mechanics, all observables are objective, i.e., they are always well defined and belong to the system or more special to the particle under observation. Moreover, anyone could measure the desired observable with any precision, without affecting the underlying system S. In other words, the properties of the systems are independent of the measurement. If one tries to do a quantum theory of measurement one has first of all to answer the question: What is the referent of quantum mechanics: measurement outcomes (the epistemic option) or object systems (the ontic, or realistic option)? In the latter case we think of QM as still describing individual particles. Here the measurement problem arises. In this branch we have furthermore to decide whether we think if quantum mechanics is universal. Do all physical systems obey the rules of quantum mechanics? If this is the case, measurement should be describable within QM. Here, serious problems immediately arise, i.e., not every observable in quantum mechanics is objective. Therefore, if the apparatus displays a value, this value need not correspond to the

observable **before** the measurement. Immediately, the question arises, how to explain and interpret the displayed value. The main idea is to define a function $p_T^E(X)$ which gives the probability of occurance of a measurement result in the set X if the system is in the state T and the observable E is measured. Those probabilities overcome the problem of objectiveness, in the sense of propensities, i.e., expressing tendencies in the behavior of individual objects. In the following we will attempt to state the basic ideas of this theory and mention the problems that occur.

5.2. Theory of the Measurement Process

5.2.1. Setting

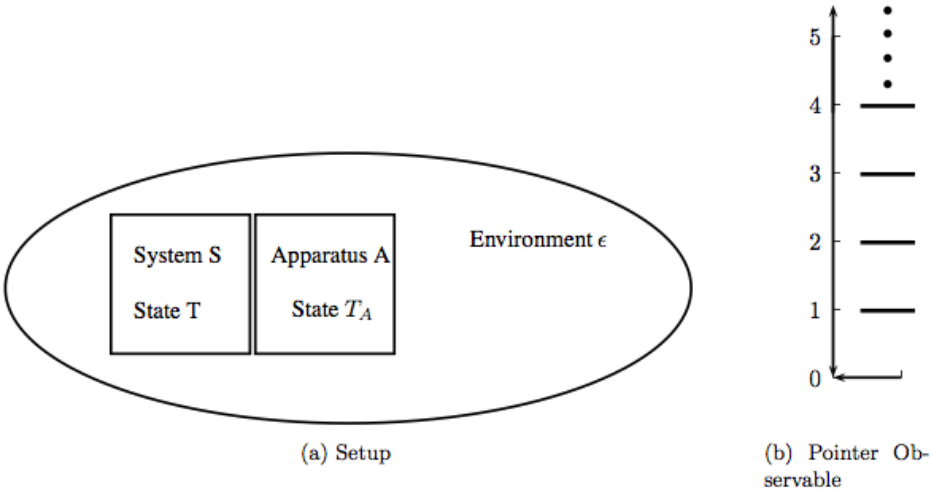

Figure 5.1: For (b) The states of the Apparatus A are macroscopically distinct. P_A therefore can have some well defined values (in this example $1, 2, 3, \ldots$)

The starting point to formulate the theory of measurement is to consider a system S in a state $T \in \mathcal{H}$. If we want to measure an observable E we have to couple the (possibly microscopical) system S to a macroscopic measuring device, denoted by apparatus A from now on. The observable one can macroscopically observe, for example some pointer on some scale is called pointer observable P_A. This pointer observable must be correlated in some

consistent way to the observable E of the system. If the P_A has some specific value, we say that E has the corresponding value.

5.2.2. Definitions

In order to work on a basis as general as possible we should define the observables (E) and states (T). These definitions which should be familiar from the last talk on projection valued measures.

Definition 8 *$\mathfrak{F}$ a σ-Algebra, $E : \mathfrak{F} \to \mathcal{L}(\mathcal{H})$ is a **projection valued measure (POV)** on $(\Omega, \mathfrak{F}) :\Leftrightarrow$*

(i) $E(X) \geq o \; \forall X \in \mathfrak{F}$

(ii) $(X_i) \in \mathfrak{F}$ countably many pairwise disjoint sets $\Rightarrow E(\cup(X_i)) = \sum_i E(X_i)$

(iii) $E(\Omega) = Id$

Note: A POV is a map from a σ-algebra $\mathfrak{F}$ of Ω to the linear operators on our Hilbert space $\mathcal{H}$

As we have seen last time the self-adjoint operators can be identified (by the spectral theorem) with the POV's. This gives a one to one correspondence between the self-adjoint operators and the POV: $A \leftrightarrow E^A$. We know from quantum mechanics that observables are always self-adjoint operators.

Definition 9 *$T \in S(\mathcal{H})$ is called a **state of the system** if it is a positive trace class operator.*

We know that for a self-adjoint operator A and a system in a pure state $\phi \in \mathcal{H}$ the measurement result of A should be $|\langle\phi| A |\phi\rangle|^2$. In our broader context we define

Definition 10 *For all E POV and all states $T \in S(\mathcal{H})$ we can define*

$$\begin{aligned} p_T^E : \quad & \mathfrak{F} \to [0,1] \\ & X \to p_T^E(X) = \mathrm{Tr}(TE(X)) \end{aligned}$$

$p_T^E(X)$ is the probability that a measurement of E performed on our System S in the state T leads to a result in X. Notice, that this coincides with what we?ve known for a long time. The language is just a bit more formal.

A short illustration of this fact comes here: Assume A to be a observable with a basis of Eigenvectors ϕ_{ij}, each eigenvalue with a $n(i)$-fold degeneracy, in other words

$$\begin{aligned} A\phi_{ij} &= a\phi_{ij} \quad \forall j = 1....n(i) \\ \text{assume } T \text{ to be a pure state } T &= P[\phi] = |\phi\rangle\langle\phi| \\ \text{clearly} \quad E^A(\{a_i\}) &= \sum_j |\phi_{ij}\rangle\langle\phi_{ij}| \end{aligned}$$

$$\begin{aligned} p_T^A(X) = \mathrm{Tr}(TA(X)) &= \sum_{a_i \in X} \mathrm{Tr}(|\phi\rangle\langle\phi| E^A(\{a_i\})) \\ &= \sum_{a_i \in X} \langle\phi| E^A(\{a_i\}) |\phi\rangle = \sum_{ij} \langle\phi|\phi_{ij}\rangle\langle\phi_{ij}|\phi\rangle \end{aligned}$$

This is clearly what you would have expected.

5.2.3. Isolated System Problem

Directly from the start occurs the so called "isolated system problem". The System and the apparatus are considered to be isolated from the environment. (Otherwise the environment would interact with the system and therefore influence the result of the measurement.) We could describe system and apparatus to be in the state T and the environment to be in the state T_ϵ. Formally the isolated compound system and environment can be described by a tensor state $W = T \otimes T_\epsilon$. Only if the time evolution is of the form $U_t = U_t^S \otimes U_T^\epsilon$ i.e., the interaction Hamiltonian between system and environment vanishes, system and environment will stay unentangled. (This is certainly not the case! More on that in the talk on decoherence) Therefore the environment does influence quantum measurements in some way.

5.2.4. Definite Outcome Problem

We know that for a given observable A we could find a basis of Eigenfunctions ϕ_{ij}. Any pure state in $\mathcal{H}$ can be written as a superposition $|\phi\rangle = \sum_I c_i |\phi_i\rangle$. Thus, in general A has no definite value. (Only if all $c_i = 0$ except one, say c_j, we could say that the system is in state $|\phi_j\rangle$). We know that our pointer observable P_A which is correlated to A does show a definite value every time we look at it. What does this value mean? This problem is called the "definite outcomes problem".

In short: The probabilistic structure of quantum mechanics gives us that even for pure states $T = p[\phi]$ clearly $p_T^E = |\langle\phi| E(X) |\phi\rangle|$ is only in special cases 0 or 1. The interpretation of this is that the system S in state T does not possess the property $E(X)$ (or its complement property). (In other words the value of the pointer is neither in X or in $\Omega - X$, respectively.) Once more we see the nonobjectivity of E.

5.2.5. Description of the Measurement Process

As above the System S is in the state $T \in S(\mathcal{H}_S)$, the apparatus A in the state $T_A \in S(\mathcal{H}_A)$. The pair of observables is denoted by E and P_A. Now we have two measure spaces, one associated with E and one with P_A respectively. Those are denoted by $(\Omega, \mathfrak{F})$ and $(\Omega_A, \mathfrak{F}_A)$. The pointer function associates values in Ω_A with those in Ω, i.e., $f : \Omega_A \to \Omega$. V is denoting the measurement coupling during the measurement.

$$V : T \otimes T_A \to V(T \otimes T_A) \tag{5.1}$$

By this coupling the states after measurement can be obtained by taking partial traces. (R takes the partial trace w.r.t. the unobserved degrees of freedom. Thus the final states are $R_S(V(T \otimes T_A))$ and $R_A(V(T \otimes T_A))$. Obviously the measurement coupling uniquely defines the final states after the measurement. This means also $p^{P_A}_{R_A(V(T\otimes T_A))}$ is completely defined. (For a given P_A).

Definition 11 *$\mathcal{M} = \langle \mathcal{H}_A P_A T_A, V, f\rangle$ is called a **measurement scheme** if the following two conditions are satisfied*

- *$\mathcal{M}$ should satisfy $p^{P_A}_{R_A(V(T\otimes T_A))}(f^{-1}(X)) = p_T^E(X)\ \forall X \in \mathcal{F}$*
- *The measurement E should yield a definite result. The pointer observable should have a definite value. "pointer objectification".*

5.3. Measurement Schemes

Definition 12 *$\mathcal{M} = \langle \mathcal{H}_A P_A T_A, V, f\rangle$ which only satisfies the first of the two properties of a measurement scheme given above is called **premeasurement** of E.*

For any observable E there exists a unitary, (V unitary) premeasurement

scheme. There have been several constructions. The first and probably easiest one was given by von-Neumann in 1932.

5.3.1. The von-Neumann pre-measurement scheme

Let O be a discrete ($\exists\Omega_0$ s.t. $O(\Omega\backslash\Omega_0) = 0$) (and sharp) observable. For O there exists a basis of eigenfunctions o_n with: (For simplicity of the argument without any degeneracies)

$$O\,|o_n\rangle = o_n\,|o_n\rangle$$

With this observable on S we associate a pointer observable P_O. This is done s.t. if the System is in the state $|o_n\rangle$ our pointer points to some specific value on a scale, lets denote this state by $|A_n\rangle$. Initially we assume the apparatus to be in the "ready state". It waits for a measurement. The system is initially prepared to be in the state $|o_n\rangle$. We know that the time-evolution in quantum mechanics is always unitary. Therefore our composite state evolves as follows:

$$|o_n\rangle \otimes |A_0\rangle \to |\tilde{o}_n\rangle \otimes |A_n\rangle$$

Now it is immediately clear, that we run in trouble, if the System is initially in a superposition state, like

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|o_l\rangle + |o_k\rangle)$$

After the measurement interaction the state evolves to:

$$|\Psi\rangle \otimes |A_0\rangle \to \frac{1}{\sqrt{2}}(|\tilde{o}_l\rangle\,|A_l\rangle + |\tilde{o}_k\rangle)\,|A_k\rangle$$

Here the trouble comes. What is the value our pointer observable should show? Should it point to l or to k ? Both results are equally likely. This is the core of the problem of "definite outcomes". This is not solved within the measurement theory. So far, assuming QM is universal and not applying any wave packed reduction postulate right here, this is an (so far) unsolved problem. Immediately the question arises: Is this measurement scheme too oversimplified? What about the environment? Maybe decoherence could make the pointer to decide to which value it should point. The short answer is: no.

5.4. The Measurement Problem Revisited

In this part of the talk we will take a look at the measurement (or macro-objectification) problem under very general assumptions. As we will see, without resorting to any of the usual assumptions of ideality, the assumption of the general validity of the linear nature of quantum mechanics while performing measurements on a microsystem leads to a fundamental contradiction.

5.4.1. Preliminaries

As we have seen in David?s talk, in the von Neumann measurement scheme for an ideal measurement process, we consider a system S, one of its observables O ($|o_n\rangle$ are the corresponding eigenvectors) and we denote by M the apparatus devised to measure O. We assume that M has a "ready-state", $|M_0\rangle$, a state in which the apparatus is ready to measure the observable and a set of *mutually orthogonal* states $|M_n\rangle$ (also orthogonal to $|M_0\rangle$ which correspond to different macroscopic configurations of the instrument.

We finally assume that the interaction between S and M is *linear* and that it yields a perfect correlation between the initial state of S and the final state of the apparatus.

$$|o_n\rangle\,|M_0\rangle \rightarrow |\tilde{o}_n\rangle\,|M_n\rangle \tag{5.2}$$

Therefore, if the final state of the apparatus is $|M_n\rangle$ then one can say that the system is in state $|o_n\rangle$.

The measurement problem arises when the system, prior to the measurement, is not in a state corresponding to one of the eigenvectors, like in (5.2), but in a superposition of them,

$$|m+n\rangle = \frac{1}{\sqrt{2}}\big(|o_m\rangle + |o_n\rangle\big) \tag{5.3}$$

In this case, the final state of the system+apparatus is given by an entangled state,

$$|m+n\rangle\otimes|M_0\rangle = \frac{1}{\sqrt{2}}\big(|o_m\rangle+|o_n\rangle\big)\otimes|M_0\rangle \rightarrow \frac{1}{\sqrt{2}}\big(|\tilde{o}_m\rangle\otimes|M_m\rangle+|\tilde{o}_n\rangle\otimes|M_n\rangle\big) \tag{5.4}$$

The essence of the quantum measurement problem is that, under the assumption that the theory is complete (i.e., that the wave-function contains

all the information about the system), it does not make sense to say that the system has any of the properties associated with $|M_m\rangle$ and $|M_n\rangle$ or in other words, the system is not in a definite macroscopic configuration.

The standard way out of this dilemma is to say that after the measurement process, the final term in (5.4) reduces to either $|\tilde{o}_m\rangle\otimes|M_m\rangle$ or $|\tilde{o}_n\rangle\otimes|M_n\rangle$ with a probability which is the square modulus of the corresponding coefficient in the superposition (in our case $1/2$). This is the *wave-packet reduction postulate* and it contradicts the general validity of the Schrödinger equation. It has been suggested that the problem derives from the over-simplified model of the von Neumann measurement scheme. In particular, the following assumptions have been criticized,

- The measuring apparatus can be prepared in a precise state $|M_0\rangle$ since the instrument is a macroscopic object with many degrees of freedom, it is impossible to know its precise state at any given time.
- One can safely neglect the interactions between the apparatus and the surrounding environment.
- The final states of the apparatus, corresponding to perceptively different macroscopic configurations of the apparatus itself, are orthogonal: actually, different states usually correspond to different positions of some component of the instrument, and since no wavefunction can have compact support in configuration space (because of the quantum evolution), wavefunctions corresponding to different states cannot, in general, be orthogonal.
- The final state of the apparatus gets perfectly correlated to the initial state of the microscopic system: this is an highly idealized characteristic which is not shared by any realistic physical instrument.

We will now address the above criticisms and show that even under very general assumptions the measurement problem cannot be avoided.

5.4.2. The Microscopic System

For simplicity we will consider the simplest system on which non-trivial measurements can be performed - a system S with a two dimensional Hilbert space, $\mathcal{H}_S$ - like the one describing the spin of an electron or the polarization

states of a photon. We consider an observable O having two different eigenvalues with $|u\rangle$ and $|d\rangle$ being the associated eigenvectors. We will refer to the property as "spin" and say that a particle has spin "up" when it is in state $|u\rangle$ and spin "down" when it is in state $|d\rangle$. We can also define superpositions of the two states, for example,

$$|u+d\rangle = \frac{1}{\sqrt{2}}(|u\rangle + |d\rangle) \tag{5.5}$$

We will assume that the system can be prepared in any one of the above considered states. We also assume that after the preparation, the system is in a precise and known state and that it can be treated as isolated from the rest of the universe, at least until the measurement process begins.

5.4.3. The Measuring Apparatus

The measuring apparatus is a macroscopic system that on interacting with the microsystem ends up in a state which is more or less correlated with the eigenstates of the observable it is devised to measure. The different outcomes of the measurement are supposed to be correlated to *perceptively different macroscopic configurations* of the apparatus - we will assume (or imagine) for simplicity that our apparatus has a pointer that moves along a scale whose position registers the result of the measurement.

Contrary to microsystems, the apparatus which is a macroscopic object has many degrees of freedom, most of which (especially the microscopic ones) we have no control over. Also, we can only have limited control over macroscopic degrees of freedom like the pointer position. Moreover, the constituents of the apparatus have existed in constant interaction with the environment for a long time before the measurement. Taking all of the above into account, we will indicate the state vectors we deal with as,

$$|A\alpha\rangle$$

Where A indicates the specific macroscopic configuration (position on the scale) of the apparatus and α refers to the condition of the rest of the universe. We will now discuss how A could be defined -

- We could say that A is the value X characterizing the projection operator $|x\rangle\langle x|$ which gives the exact position of the center of mass of the

pointer. However, this is not possible as it is impossible to measure a continuous variable with perfect accuracy.

- A possible improvement is saying that "the pointer is at position x" when the wave-function is an eigenstate of the projection operator which projects onto the interval $\Delta(x) = [x - \delta, x + \delta]$. However, this is also not possible as Hamiltonian evolution transforms any wave-function with compact support into one with non-compact support

In order to overcome the above problems, we consider a very general physical system - we define the set V_A as the set of all normalized vectors $|A\alpha\rangle$ for which we are *allowed* to say that the pointer of the apparatus is in position A. There is no restriction on the type of wave-functions that belong to V_A. The only physical restriction we put forward on the system is that if the pointer admits two physically distinct positions (say A and B) then any two vectors corresponding to these different positions should be *almost orthogonal*. We can express this requirement mathematically as,

$$\inf_{|A\alpha\rangle \epsilon V_A, |B\beta\rangle \epsilon V_B} \| \, |A\alpha\rangle - |B\beta\rangle \, \| \geq \sqrt{2} - \eta \ , \ \ \eta \ll 1 \tag{5.6}$$

Here $|A\alpha\rangle$ and $|B\beta\rangle$ are normalized to 1. The above requirement is necessary if we want that *different* macroscopic configurations of the pointer represent *mutually exclusive* configurations of the microsystem.

The index α takes into account all the degrees of freedom that are not in our control. Two vectors labelled by A but with different values for α, refer to the "same" macroscopic configuration for our pointer but describe two different states for the rest of the universe. For example, a certain atom of the apparatus might be in the ground state for $|A\alpha\rangle$ but might be in an excited state for $|A\beta\rangle$.

Since our microsystem has two spin states and we want the apparatus to distinguish between them, we must assume that our pointer admits two macroscopically distinct positions, U and D along the scale; in light of the previous argument we define two sets V_U and V_D such that the first contains all the (normalized) vectors for which we are allowed to say that the system is in a state "up" and similarly the second for the state "down". We define two vectors chosen from the two different sets to be almost orthogonal in the

sense of (5.6), i.e.,

$$\inf_{|U\alpha\rangle\epsilon V_U, |D\beta\rangle\epsilon V_D} \| \, |U\alpha\rangle - |D\beta\rangle \, \| \geq \sqrt{2} - \eta \ , \ \ \eta \ll 1 \tag{5.7}$$

One property of the sets V_U and V_D which we can immediately see is that they contain no vectors in common. If this were the case, then the distance between any two such vectors would be zero and this would contradict (5.7).

5.4.4. Preparation Of The Apparatus

We denote the ready state of the apparatus as A_0. The initial state of the apparatus is given by,

$$|A_0\alpha\rangle = |spin\rangle \otimes |A_0\bar{\alpha}\rangle \tag{5.8}$$

where $\bar{\alpha}$ refers to the state of the universe with the exception of the pointer and the particle (microsystem) and $|spin\rangle$ refers to the initial state-vector of the particle.

In general, we have no control over $\bar{\alpha}$, so we do not know the precise initial state. We can say that any specific $\bar{\alpha}$ occurs with a probability $p(\bar{\alpha})$ which is also unknown to us. We will later discuss what requirements this probability must satisfy. For the moment, the inital setup of the whole particle+apparatus+universe system will be described as,

$$\{|spin\rangle \otimes |A_0\bar{\alpha}\rangle \, , p(\bar{\alpha})\} \tag{5.9}$$

5.4.5. The Measurement Process

If we assume that quantum mechanics governs all physical systems then the measurement process, being an interaction between two quantum systems is governed by a unitary operator $U(t_I, t_F)$ for an interaction that occurs between the times t_I and t_F. Suppose that the initial state of the microsystem is $|u\rangle$ and the apparatus is in the ready-state, then we can say that during the measurement process, the whole universe evolves in the following way,

$$|u\rangle \otimes |A_0\bar{\alpha}\rangle \rightarrow U(t_I, t_F)[|u\rangle \otimes |A_0\bar{\alpha}\rangle] = |Fu\bar{\alpha}\rangle \tag{5.10}$$

If the initial state of the microsystem it is,

$$|d\rangle \otimes |A_0\bar{\alpha}\rangle \rightarrow U(t_I, t_F)[|d\rangle \otimes |A_0\bar{\alpha}\rangle] = |Fd\bar{\alpha}\rangle \tag{5.11}$$

- Since the evolution is unitary, once the initial state is specified, the final state is unambiguously determined. Accordingly the final state is also characterized by the same index $\bar{\alpha}$.

- Unlike the von Neumann measurement scheme, the final state is not factorized,
$$|Fu\bar{\alpha}\rangle \neq |u\rangle \otimes |A_U\bar{\alpha}\rangle \tag{5.12}$$

- We allow for the fact that the measurement process can affect the particle in a non-trivial way. The final state of the microsystem isn't the same as the initial state.

The only requirement we put forward is that the measuring apparatus should be reliable to a *high degree*. This means that if the initial state is $|u\rangle$ then the final state $|Fu\bar{\alpha}\rangle$ should belong to V_U in most of the cases and likewise for an initial state $|u\rangle$. We can formalize the reliability requirements in the following way - Consider the set K of all subsets J of the possible values that $\bar{\alpha}$ can assume such that J is equipped with the following natural measure,

$$\mu(J) = \sum_{\bar{\alpha}\epsilon J} p(\bar{\alpha}) \tag{5.13}$$

And we define the following two sets,

$$J_U^- = \{\bar{\alpha} : |Fu\bar{\alpha}\rangle \notin V_U\} \tag{5.14}$$
$$J_D^- = \{\bar{\alpha} : |Fd\bar{\alpha}\rangle \notin V_D\} \tag{5.15}$$

$J_{U/D}^-$ are the set of indices that correspond to the result that the pointer is not in the position U/D after measurement despite the fact that the initial state of the microsystem was $|u/d\rangle$. We also define J_U^+ as the complement of J_U^- and J_D^+ as the complement of J_D^-. The reliability requirement can be expressed in the following way - given $\epsilon \ll 1$

$$\mu(J_U^-) \leq \epsilon \qquad \mu(J_D^-) \leq \epsilon \tag{5.16}$$

The parameter ϵ expresses the efficiency of the measuring apparatus and can depend on the actual measurement process devised. It is easy to see that the complement sets satisfy the following relations,

$$\mu(J_U^+) \geq 1 - \epsilon \qquad \mu(J_D^+) \geq 1 - \epsilon \tag{5.17}$$

We should also take into account the following sets, $J^- = J_U^- \cup J_D^-$ and its complement $J^+ = J_U^+ \cap J_D^+$. They satisfy the following relations,

$$\mu(J^-) \leq 2\epsilon \qquad \mu(J^+) \geq 1 - 2\epsilon \tag{5.18}$$

While we have taken into account the possibility of a wrong measurement, we have also taken into account the case when the apparatus fails to interact with the microsystem. In this case the vector belongs to J^-.

In the case that the apparatus fails to detect the particle, there can be an appreciable affect in many experimental situations (e.g., the efficiency of photodetectors is usually quite low). We can easily circumvent this difficulty by simply disregarding all cases in which a detector should register something but it doesn't. The previous analysis should therefore be interpreted as referring only to cases where the apparatus registers an outcome.

5.4.6. Result and Conclusions

As stated earlier, the measurement problem arises when the microsystem is prepared in a superposition state which is not the same as the individual $|u\rangle$ or $|d\rangle$ states. With such a state, due to the linearity of quantum evolution, the resulting final state of the particle+apparatus system will be,

$$\begin{aligned} |u+d\rangle \otimes |A_0\bar{\alpha}\rangle &\rightarrow U(t_I, t_F)[|u+d\rangle \otimes |A_0\bar{\alpha}\rangle] \\ &= |F(u+d)\bar{\alpha}\rangle \\ &= \frac{1}{\sqrt{2}}[|Fu\bar{\alpha}\rangle + |Fd\bar{\alpha}\rangle] \end{aligned} \tag{5.19}$$

We will now prove that for each $\bar{\alpha}$ belonging to J^+, $|F(u+d)\bar{\alpha}\rangle$ cannot belong to either V_U or V_D. Suppose that $|F(u+d)\bar{\alpha}\rangle \in V_U$ (the proof for V_D is analogous). The distance between $|F(u+d)\bar{\alpha}\rangle$ and $|Fd\bar{\alpha}\rangle$ is,

$$\| \, |F(u+d)\bar{\alpha}\rangle - |Fd\bar{\alpha}\rangle \, \| = \| 1/\sqrt{2} \, |Fu\bar{\alpha}\rangle + (1/\sqrt{2} - 1) \, |Fd\bar{\alpha}\rangle \, \| \leq 1 \tag{5.20}$$

Which contradicts our "almost orthogonal" requirement shown in (5.7) since $|F(u+d)\bar{\alpha}\rangle \in V_U$ and $|Fd\bar{\alpha}\rangle \in V_D$. This shows that for $\bar{\alpha} \in J^+$, $|F(u+d)\bar{\alpha}\rangle \notin V_U \text{ or } V_D$.

We have therefore seen that for all $\bar{\alpha} \in J^+$ and for all measurement processes for which the apparatus registers an outcome, the vector $|F(u+d)\bar{\alpha}\rangle$ does not allow us to assign a definite macroscopic position to the pointer of the apparatus.

6. EPR, Bell's inequality and the Kochen-Specker theorem - Isabel Krebs and Matthias Schlaffer

6.1. Introduction

These are the notes to our presentation held in the Seminar "Foundations of quantum mechanics". When reading, be aware of possible faults and typos.

When quantum mechanics was developed in the beginning of the 20^{th} century, Al- bert Einstein, Boris Podolsky and Nathan Rosen wrote their famous paper about the completeness of quantum theory [9]. In this paper they argued that quantum mechanics cannot be complete theory. However this argument was based on the assumption of locality and realism.

Later, in 1964 John S. Bell reformulated this idea in mathematical terms, developing an inequality which must hold for any local and realistic theory. He showed that this inequality was violated by the predictions of quantum theory [3].

Finally Simon B. Kochen and Ernst Specker proved in 1986 that no realistic and non-contextual theory can reproduce quantum mechanical predictions [14]. Since the original proof is rather complicated and not very instructive, we will present the much easier version formulated by Nathaniel D. Mermin in 1990 [16].

6.2. Definitions

Locality: Systems that are separated by a space-like distance cannot influence each other.

Realism: Every observable of a physical system has a definite value at any time.

Non-contextuality: The result of a measurement of an observable A does not depend on the measurement context, i.e., it does not depend on whether A is simultaneously measured with B and C or D and E where $\{A, B, C\}$ and $\{A, D, E\}$ are two sets of pairwise commuting observables (the observables B and D do not commute in general).

Complete theory: A theory is complete if for every element of reality there is a corresponding element in the theory.

6.3. Einstein-Podolsky-Rosen argument

The EPR argument shows that, assuming locality and realism one can conclude that quantum mechanics is not a complete theory. In order to get to this conclusion, they considered a gedanken experiment based on quantum mechanics which uses a two particle entangled state.

While the observables in the original version were the positions and momenta of the two particles, modern versions use two components of the spin of two spin-1/2 particles in a singlet state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$$

After interacting, the two particles are separated by a space-like distance. The x- or y-component of the spin (σ_x and σ_y respectively) of each particle can be measured by an apparatus. When measuring σ_x of the first particle the result is either +1 or −1. If for example the result is +1 a measurement of the σ_x of the second particle will always give the result −1 and vice versa. So by performing a measurement on one particle information about the other particle can be gained.

Since one can decide after the space-like separation of the particles whether to measure the x- or y-component of the spin and locality is assumed one has to conclude that the values of both components of the spin of both particles must be predetermined.

In the EPR paper realism is defined as follows: "*If, without in any way disturbing a system, we can predict with certainty [...] the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.*" [9]. According to this definition, the components of both spins of the previously mentioned gedanken experiment have simultaneously physical reality. However quantum mechanics states that if two observables do not commute, it is impossible to know simultaneously the value of both of them. EPR conclude that the quantum mechanical description is incomplete because it does not provide information about these values.

The remaining question is if quantum mechanics can be completed to a realistic theory, i.e., can we introduce so called hidden variables by which the outcome of a measurement can be predicted with certainty? Bell's theorem provides one restriction to these hidden variable theories: No local hidden variable theory can reproduce the predictions of quantum mechanics.

6.4. The original Bell argument

By considering realism and locality, Bell achieved to derive an inequality. Several experiments proved this inequality to be wrong which means that nature is not local *and* realistic.

To derive the inequality consider two space-like separated spin-1/2 particles in the singlet state. The observable that measures the spin of the first particle is given by $\sigma_1 \cdot \mathbf{a}$ where $\mathbf{a}$ is a unit vector determining the direction of the spin measurement. Similarly the observable $\sigma_2 \cdot \mathbf{b}$ measures the spin of the second particle in direction of the unit vector $\mathbf{a}$. Since realism is assumed the result of the measurement of $\sigma_1 \cdot \mathbf{a}$ can be written as a function of $\mathbf{a}$ and a hidden variable λ: $A(\mathbf{a}, \lambda) = \pm 1$, likewise the result of $\sigma_2 \cdot \mathbf{b}$ is given by $B(\mathbf{b}, \lambda) = \pm 1$. It is not necessary to specify λ, it could be just a number or a vector but also a more complex object like a function. Note that since locality is assumed and the particles are space-like separated A cannot depend on $\mathbf{b}$ and B not on $\mathbf{a}$. The expectation value $E(\mathbf{a},\mathbf{b})$ of the product of $\sigma_1 \cdot \mathbf{a}$ and $\sigma_2 \cdot \mathbf{b}$ is therefore given by:

$$E(\mathbf{a}, \mathbf{b}) = \int d\lambda \, \rho(\lambda) A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda) \tag{6.1}$$

where $\rho(\lambda)d\lambda$ is the probability measure of λ obeying $\int d\lambda \, \rho(\lambda) = 1$ and $\rho(\lambda) \geq 0 \; \forall \lambda$. The quantum mechanical expression for this expectation value is given by

$$\langle \sigma_1 \mathbf{a} \cdot \sigma_2 \mathbf{b} \rangle = -\mathbf{a} \cdot \mathbf{b} = -\cos\phi \tag{6.2}$$

with ϕ being the angle between $\mathbf{a}$ and $\mathbf{b}$.

Of course the two expressions (6.1) and (6.2) should be equal for all $\mathbf{a}$ and $\mathbf{b}$, in particular for $\mathbf{a} = \mathbf{b}$ where quantum mechanics predicts an expectation value of -1. Since $A(\mathbf{a}, \lambda) = \pm 1$, $B(\mathbf{b}, \lambda) = \pm 1$ and $\int d\lambda \, \rho(\lambda) = 1$ (6.1) can achieve this value only if $A(\mathbf{a}, \lambda) = -B(\mathbf{b}, \lambda) \; \forall \lambda : \rho(\lambda) \neq 0$. If for some λ with

$\rho(\lambda) \neq 0$ the product $A(\mathbf{a},\lambda)B(\mathbf{b}$ is $+1$ there will be a positive contribution to the integral and thus $E(\mathbf{a},\mathbf{b}) > -1$. So (6.1) can be rewritten as:

$$E(\mathbf{a},\mathbf{b}) = \int d\lambda\, \rho(\lambda) A(\mathbf{a},\lambda) A(\mathbf{b},\lambda) \tag{6.3}$$

Now let $\mathbf{c}$ be another unit vector, then

$$\begin{aligned} E(\mathbf{a},\mathbf{b}) - E(\mathbf{a},\mathbf{c}) &= -\int d\lambda\, \rho(\lambda)[A(\mathbf{a},\lambda)A(\mathbf{b},\lambda) - A(\mathbf{a},\lambda)A(\mathbf{c},\lambda)] \\ &= \int d\lambda\, |\rho(\lambda)[A(\mathbf{a},\lambda)A(\mathbf{c},\lambda) \underbrace{A^2(\mathbf{a},\lambda)}_{=1} - A(\mathbf{a},\lambda)A(\mathbf{b},\lambda)]| \\ &= \int d\lambda\, \rho(\lambda) A(\mathbf{a},\lambda)A(\mathbf{b},\lambda)[A(\mathbf{b},\lambda)A(\mathbf{c},\lambda) - 1] \end{aligned} \tag{6.4}$$

Finally apply the absolute value to this equation:

$$\begin{aligned} |E(\mathbf{a},\mathbf{b}) - E(\mathbf{a},\mathbf{c})| &= \left| \int d\lambda\, \rho(\lambda) A(\mathbf{a},\lambda)A(\mathbf{b},\lambda)[A(\mathbf{b},\lambda)A(\mathbf{c},\lambda) - 1] \right| \\ &\leq \int d\lambda\, |\rho(\lambda) \underbrace{A(\mathbf{a},\lambda)A(\mathbf{b},\lambda)}_{=\pm 1} [A(\mathbf{b},\lambda)A(\mathbf{c},\lambda) - 1]| \\ &\leq \int d\lambda\, \rho(\lambda)[1 - A(\mathbf{b},\lambda)A(\mathbf{c},\lambda)] \\ &= 1 + E(\mathbf{b},\mathbf{c}) \end{aligned} \tag{6.5}$$

The result is the famous Bell inequality:

$$1 + E(\mathbf{b},\mathbf{c}) \geq |E(\mathbf{a},\mathbf{b}) - E(\mathbf{a},\mathbf{c})| \tag{6.6}$$

It can be easily seen that quantum mechanical predictions contradict this inequality. Choose for example

$$\mathbf{a} = \frac{1}{\sqrt{2}}(-\hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y) \quad \mathbf{b} = \hat{\mathbf{e}}_y \quad \mathbf{c} = \hat{\mathbf{e}}_x \tag{6.7}$$

then

$$E(\mathbf{b},\mathbf{c}) = 0 \quad E(\mathbf{a},\mathbf{b}) = -E(\mathbf{a},\mathbf{c}) = \frac{1}{\sqrt{2}} \tag{6.8}$$

and the Bell inequality states

$$1 \geq \sqrt{2} \tag{6.9}$$

which is clearly wrong.

The experimental results verify quantum mechanical predictions and thus violate the Bell inequality. Therefore there cannot be a local and realistic theory describing nature. The only problem about the Bell inequality is that statistical values are used. In general it is impossible to measure an expectation value. It can only be achieved by averaging over many measurements. There are, however, similar arguments circumventing this problem as it will be shown in section 6.6.

6.5. The Kochen-Specker theorem

The Kochen-Specker theorem states that there is no realistic and non-contextual theory describing nature:

Theorem 13 *It is, in general, impossible to find a map ν which maps the operators $A, B, C, ...$ to their measurement result $\nu(A), \nu(B), \nu(C), ...$ and*

1. *$\nu(X)$ is an eigenvalue of $X \in \{A, B, C, ...\}$*
2. *If $A, B, C, ...$ are mutually commuting and $f(A, B, C, ...) = 0$, then $f(\nu(A), \nu(B), \nu(C), ...) = 0$*

The map ν states the realism since it allows, at least theoretically, to predict the result of a measurement with certainty. The first condition is dictated by quantum mechanics since a measurement result is always an eigenvalue of the measured operator. The second condition is the mathematical expression for non-contextuality, stating that the result of measuring A is independent of the other operators and can therefore be considered separately. In the case where f just multiplies the operators this leads to the conclusion that $\nu(ABC\cdots) = \nu(A)\nu(B)\nu(C)\cdots$,

Proof. Consider 3 spin-½ particles in an arbitrary state and a map obeying the assumptions of theorem 13. Since the eigenvalue of $\sigma = \pm 1$ this leads to the following equations:

$$\nu(\sigma_x^1\sigma_y^2\sigma_y^3)\nu(\sigma_x^1)\nu(\sigma_y^2)\nu(\sigma_y^3) = 1 \tag{6.10a}$$

$$\nu(\sigma_y^1\sigma_x^2\sigma_y^3)\nu(\sigma_y^1)\nu(\sigma_x^2)\nu(\sigma_y^3) = 1 \tag{6.10b}$$

$$\nu(\sigma_y^1\sigma_y^2\sigma_x^3)\nu(\sigma_y^1)\nu(\sigma_y^2)\nu(\sigma_x^3) = 1 \tag{6.10c}$$

$$\nu(\sigma_x^1\sigma_y^2\sigma_y^3)\nu(\sigma_y^1\sigma_x^2\sigma_y^3)\nu(\sigma_y^1\sigma_y^2\sigma_x^3)\nu(\sigma_x^1\sigma_x^2\sigma_x^3) = -1 \tag{6.10d}$$

As the sigma matrices belonging to different particles commute, the first factor in equations (6.10a)-(6.10c) can be factorized as described above. Then each factor of these equations appears twice, i.e. squared, and thus they are all equal to 1. The expression in equation (6.10d) is equal to −1 since $(\sigma_y^1\sigma_y^2\sigma_x^3)\cdot(\sigma_y^1\sigma_x^2\sigma_y^3)\cdot(\sigma_y^1\sigma_y^2\sigma_x^3) = -(\sigma_x^1\sigma_x^2\sigma_x^3)$, which can be calculated by the commutator relations of sigma matrices.

When multiplying the left hand sides of equations (6.10a)-(6.10d) one gets +1 since each factor appears squared, multiplying the right hand sides however yields −1, a contradiction.■ The advantage of the Kochen-Specker argument over Bell's argument is that for its proof no specific state needed to be considered while for Bell's argument a singlet state had to be assumed.

6.6. Another version of Bell's theorem

The operators of the above presented Kochen-Specker theorem can be also used for a Bell argument which, in contrast to the original one, has no statistical character. To see this, consider the system to be in a state Φ being an eigenstate of the mutually commuting operators $\sigma_x^1\sigma_y^2\sigma_y^3, \sigma_y^1\sigma_x^2\sigma_y^3$ and $\sigma_y^1\sigma_y^2\sigma_y^3$ with eigenvalue +1. Since $(\sigma_y^1\sigma_y^2\sigma_x^3)\cdot(\sigma_y^1\sigma_x^2\sigma_y^3)\cdot(\sigma_y^1\sigma_y^2\sigma_x^3) = -(\sigma_x^1\sigma_x^2\sigma_x^3)$, Φ is also an eigenstate of $(\sigma_x^1\sigma_x^2\sigma_x^3)$ with eigenvalue −1.

Suppose now three space-like separated particles in this state Φ. It is possible to find out the result m_x of a spin measurement in x-direction of one particle by measuring the y-components of the spin of the other two particles, since the product of these three values must be 1. Similarly one can obtain the result of a spin measurement in y-direction by measuring the spin of the other two particles in x- and y-direction, respectively.

The assumption of locality and realism leads to the conclusion, that the system in state Φ must be describable by the six variables $m_x^1, m_y^1, m_x^2, m_y^2, m_x^3$ and m_y^2 which determine the result of the measurements of the above mentioned observables. Also since Φ is an eigenstate of these observables the m must fulfill these equations:

$$m_x^1 m_y^2 m_y^3 = 1 \tag{6.11}$$

$$m_y^1 m_x^2 m_y^3 = 1 \tag{6.12}$$

$$m_y^1 m_y^2 m_x^3 = 1 \tag{6.13}$$

$$m_x^1 m_x^2 m_x^3 = -1 \tag{6.14}$$

Now as in section 13 the product of the left hand sides of equations (6.11)-(6.14) yields +1 since every number appears twice, i.e. squared, while the product of the right hand sides yields −1. This is again a contradiction confirming the result of the original Bell argument that a quantum theory cannot be local and realistic.

6.7. Conclusion

The theorems of Bell and Kochen-Specker are the two most important so-called no-go theorems. They constrain the possibilities of a quantum theory. Bell?s theorem states that a quantum theory cannot be both local and realistic. This has been confirmed by experiments. The Kochen-Specker theorem forbids theories which are realistic as well as non-contextual.

Therefore a quantum theory describing nature properly must sacrifice either locality and non-contextuality or realism or all these three concepts. Since these theorems do not specify which of these concepts are to be given up, there are several theories. While the standard literature tends to abandon either both concepts or the concept of realism, bohmian mechanics for example sticks to realism and gives up locality and non-contextuality.

7. Decoherence - Kostas Vavouranakis and Kosmas Kepesidis

7.1. Abstract

Decoherence is neither an interpretation nor a modification of quantum mechanics. On the contrary, it exists within the quantum formalism. The decoherence program attempts to answer the question “How can a world, being fundamentally quantum in nature, appear to be classical?”

7.2. Decoherence

7.2.1. Resolution into subsystems

Let us start with some obvious facts. First, that the whole universe is by definition a closed system. And second, that a closed system is described by Schrödinger's equation (unitary time evolution etc).

Decoherence comes from the possibility to divide the total system into (sub)systems and environment. This openess of quantum systems, which is an essential requirement for decoherence to occur, is formally described by the decomposition of the total Hilbert space into a tensor product of the Hilbert spaces of the subsystems and the total state into reduced states.

$$\mathcal{H}_{total} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_N \tag{7.1}$$

$$|\psi_{total}\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_N\rangle \tag{7.2}$$

But how exactly can one define the environment?

Definition: The environment consists of all those degrees of freedom that cannot be controlled and their dynamics is in general unknown.

7.2.2. The Reduced Density Matrix

Consider a pair of spins being initially in an entangled state. For instance, a state of the EPR type,

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2) \tag{7.3}$$

The density matrix of the total system is,

$$\rho = |\psi\rangle \langle\psi| \tag{7.4}$$

Now, consider an observable that refers only to the first of those two spins,

$$\hat{O} = \hat{O}_i \otimes \hat{I}_2 \tag{7.5}$$

then the expectation value of the observable reads,

$$\langle\hat{O}\rangle_\psi = \mathrm{Tr}(\rho\hat{O}) = \mathrm{Tr}_1(\hat{O}_1 \mathrm{Tr}_2(\rho)) = \mathrm{Tr}_1(\hat{O}_1\rho_1) \tag{7.6}$$

where $\rho_1 = \mathrm{Tr}_2(\rho)$ is the reduced density matrix (RDM) of the first spin. It constitutes a local representation since the degrees of freedom of the second spin have been traced out. Hence the RDM reads,

$$\rho_1 = \mathrm{Tr}_2 |\psi\rangle\langle\psi| = {}_2\langle + |\psi\rangle\langle\psi|+\rangle_2 + {}_2\langle - |\psi\rangle\langle\psi|-\rangle_2 = \frac{1}{2}(|+\rangle\langle +|)_1 + \frac{1}{2}(|-\rangle\langle -|)_1 \tag{7.7}$$

Where we have used the fact that the spin states are well distinguishable, i.e., mutually orthogonal ${}_2\langle + |-\rangle_2 = 0$. It is obvious from equation (7.7) that the RDM has diagonal form. Thus an important result is that the amount of overlap of the states of system 2 that are one-to-one correlated with the states of system 1, quantifies the degree of interference in the basis $[|+\rangle, |-\rangle]$ of system 2. For that the diagonalization of the RDM is basis dependent. In other words the wavefunction (7.3) corresponds to a maximally entangled state. On the other extreme, a tensor product state of the general form

$$|\psi'\rangle = \frac{1}{\sqrt{2}}(|+\rangle_1 + e^{i\phi}|-\rangle_1)|\Phi\rangle_2$$

where there are no correlations between the states of the system 1 and those of the system 2, does not lead to a diagonal RDM.

A diagonal density matrix represents a classical statistical mixture. That means that in this case the outcome of a measurement on the spin 1 would be + with probability $\frac{1}{2}$ and − with probability $\frac{1}{2}$. Those probabilities are classical and there is a clear analogy to the case that we flip a coin. Classical probabilities are always related to ignorance (lack of information). For example coin flipping is a completely deterministic process, however the outcome can only be described by probabilities because a complete description is impossible due to the complexity of the problem. That means we ignore details. Similarly, in the case of the EPR state, the details we ignore are the quantum correlations between the two spins.

7.2.3. General formalism

In section 2.2 we saw that the diagonalization of the RDM was implied by the fact that the spin states are well distinguishable. Now, we will calculate the RDM in the general case of a quantum system interacting with its environment. Suppose a quantum system being initially in a pure state

$$|\psi\rangle = \sum_n c_n |s_n\rangle$$

where $[|s_n\rangle]$ are basis vectors in a Hilbert space $\mathcal{H}_s$ and its environment, which can also be a quantum system

$$|\mathcal{E}\rangle = \sum_n d_n |e_n\rangle$$

with vectors $[|e_n\rangle]$ in $\mathcal{H}_e$ which in general are not orthogonal. We suppose that initially the system and the environment are uncorrelated.

$$|\psi_{total}\rangle = |psi\rangle \otimes |\mathcal{E}\rangle \tag{7.8}$$

Now, if we turn on interactions between them, according to the von Neumann scheme, after unitary time evolution the total system will end up in an entangled state,

$$|\psi_{total}\rangle \to \left|\tilde{\psi}_{total}\right\rangle = \sum_{n,m} c_{nm} |s_n\rangle |e_m\rangle = \sum_n \tilde{c}_n |\tilde{s}_n\rangle |\tilde{e}_n\rangle \tag{7.9}$$

For the last step in the above equation we have used the Schmidt decomposition theorem, which states that an arbitrary pure state of a bipartite system can always be written in a diagonal form. Therefore, the density matrix of the combined system reads,

$$\rho_{se} = \sum_{mn} \tilde{c}_m \tilde{c}_n^* |\tilde{s}_m\rangle |\tilde{e}_m\rangle \langle\tilde{s}_n| \langle\tilde{e}_n| \tag{7.10}$$

And by taking the trace over the degrees of freedom of the environment we get the RDM for the system,

$$\rho_s = \mathrm{Tr}_e(\rho_{se}) = \sum_{mn} \tilde{c}_m \tilde{c}_n^* |\tilde{s}_m\rangle \langle\tilde{s}_n| \langle\tilde{e}_n | \tilde{e}_n\rangle \tag{7.11}$$

Up to now, the RDM contains interference terms (coherences) for $n \neq m$, since, in general, the vectors of the environment are not mutually orthogonal. However, many physical simulations have shown that due to the large number of subsystems that com- pose the environment, the states of the environment rapidly approach orthogonality $\langle\tilde{e}_n | \tilde{e}_n\rangle (t) \to \delta_{n,m}$. This leads to the loss of the coherence terms and it will be shown in the next section using a generic example.

$$\rho_s \to \tilde{\rho}_s = \sum_n |\tilde{c}_n|^2 |\tilde{s}_m\rangle \langle\tilde{s}_n| \tag{7.12}$$

We must also mention that, in the case of a measurement, the orthogonality of the states of the environment is a physical requirement for the effectiveness of the measurement apparatus, i.e., the ability to give well specified outcomes. The phenomenon of the dynamical diagonalization of the RDM is called Environment-Induced-Decoherence.

7.2.4. An example on decoherence

To illustrate how decoherence takes place in practice let us consider a realistic and rather general example. Consider the world composing out of 2-state systems. We shall investigate one of them. In practice we have a 2-state system interacting with N 2-state "environments". The Hamiltonian for the interaction is of the following form

$$\begin{aligned} H_{se} =& (|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|) \\ & \otimes \sum_k g_k (|\uparrow_k\rangle\langle\uparrow_k| - |\downarrow_k\rangle\langle\downarrow_k|) \bigotimes_{k\neq k'} \underbrace{(|\uparrow_k\rangle\langle\uparrow_k| + |\downarrow_k\rangle\langle\downarrow_k|)}_{I_{k'}} \end{aligned} \tag{7.13}$$

with I_k being the identity operator for the k^{th} environmental spin and g_k coupling constants.

Expanding the $|\uparrow\rangle$ and $|\downarrow\rangle$ basis the initial state is

$$|\psi(0)\rangle = (a|\uparrow\rangle + b|\downarrow\rangle) \bigotimes_{k=1}^{N} (\alpha_k|\uparrow_k\rangle + \beta_k|\downarrow_k\rangle) \tag{7.14}$$

And after the interaction is turned on, by solving Schrödinger's equation we get

$$\begin{aligned} |\psi(t)\rangle = a|\uparrow\rangle & \underbrace{\bigotimes_{k=1}^{N} (\alpha_k e^{ig_k t}|\uparrow_k\rangle + \beta_k e^{-ig_k t}|\downarrow_k\rangle)}_{\equiv |E_\uparrow(t)\rangle} \\ + b|\downarrow\rangle & \underbrace{\bigotimes_{k=1}^{N} (\alpha_k e^{-ig_k t}|\uparrow_k\rangle + \beta_k e^{ig_k t}|\downarrow_k\rangle)}_{\equiv |E_\downarrow(t)\rangle} \end{aligned}$$

Therefore we have the evolution of our initial state into an entangled state

$$|\psi t\rangle = a|\uparrow\rangle|E_\uparrow(t)\rangle + b|\downarrow\rangle|E_\downarrow(t)\rangle \tag{7.15}$$

The density matrix rea

$$\begin{aligned} \rho = |\psi(t)\rangle\langle\psi(t)| = & |a|^2|\uparrow\rangle\langle\uparrow| \otimes |E_\uparrow(t)\rangle\langle E_\uparrow(t)| \\ & + |b|^2|\downarrow\rangle\langle\downarrow| \otimes |E_\downarrow(t)\rangle\langle E_\downarrow(t)| \\ & + ab^*|\uparrow\rangle\langle\downarrow| \otimes |E_\uparrow(t)\rangle\langle E_\downarrow(t)| \\ & + a^*b|\downarrow\rangle\langle\uparrow| \otimes |E_\downarrow(t)\rangle\langle E_\uparrow(t)| \end{aligned}$$

Now we can estimate the reduced density matrix by tracing out the environment

$$\rho_s = |a|^2 |\uparrow\rangle \langle\uparrow| + |b|^2 |\downarrow\rangle \langle\downarrow| + z(t) ab^* |\uparrow\rangle \langle\downarrow| + z^*(t) a^* b |\downarrow\rangle \langle\uparrow| \tag{7.16}$$

where $z(t) = \langle E_\downarrow(t) \,|\, E_\uparrow(t)\rangle$ are the coefficients for the interference terms. Let us study how these coefficients evolve in time. Using the explicit form of the states $|E_\downarrow(t)\rangle$ and $|E_\uparrow(t)\rangle$ we get

$$z(t) = \prod_{k=1}^{N} (|\alpha_k|^2 e^{2ig_k t} + |\beta_k|^2 e^{-2ig_k t}) \tag{7.17}$$

For simplicity, we assume that all g_k's, $|\alpha_k|$'s and $|\beta_k|$'s have no dependence on the index k, namely that $g_k = g$, $|\alpha_k| = \alpha$ and $|\beta_k| = \beta$ (alternatively, we could assume they are drawn from some probability distribution and then appeal to the central limit theorem to obtain the Gaussian as below). Then

$$z(t) = (|\alpha|^2 e^{2igt} + |\beta|^2 e^{-2igt})^N = \sum_{l=0}^{N} \binom{N}{l} |\alpha|^{2l} |\beta|^{2(N-l)} e^{2ig(2l-N)t} \tag{7.18}$$

which is a binomial distribution. In the large N limit the binomial distribution can be approximated by a Gaussian.

$$\binom{N}{l} |\alpha|^{2l} |\beta|^{2(N-l)} \approx \frac{1}{\sqrt{2\pi N |\alpha|^2 |\beta|^2}} \exp\left(\frac{-(l - N|\alpha|^2)^2}{2N|\alpha|^2|\beta|^2} \right)$$

Given this approximation equation (7.18) can be written as

$$z(t) = \sum_{l=0}^{N} \frac{1}{\sqrt{2\pi N |\alpha|^2 |\beta|^2}} \exp\left(\frac{-(l - N|\alpha|^2)^2}{2N|\alpha|^2|\beta|^2} \right) e^{2ig(2l-N)t} \tag{7.19}$$

The above is reminiscent of the Fourier transform of a Gaussian, which should be a Gaussian itself. Let us bring the above equation into a more suggestive form, so that the Gaussian and the Fourier transform along with the conjugate variable become apparent. To simplify the notation, we define $A = (2N|\alpha|^2|\beta|^2)^{-1}$ and we make a variable substitution $k = l - N|\alpha|^2$. Our previous claim is now obvious

$$z(t) = \frac{A}{\sqrt{\pi}} e^{-2igN(1-|\alpha|^2)t} \sum_k e^{-Ak^2} e^{4igkt} \tag{7.20}$$

By treating the above in the usual way, after completing the square of the exponent, substituting back the explicit form of the quantity A and replacing $\sum_k$ by $\int dk$ we get the final result

$$z(t) \approx \frac{\exp\left(-2igN(1-|\alpha|^2 t)\right.}{\sqrt{2\pi N}|\alpha||\beta|} \exp\left(-4g^2\sqrt{2N}|\alpha||\beta|t^2\right) \tag{7.21}$$

We are now in a position to comment on the quantum to classical transition of the system under study. Analyzing the formula for the coefficients functions of the ???interference, we see that it is composed out of an oscillating factor (which is of no particular interest for our analysis), one suppressing factor $(1/\sqrt{N})$ and a Gaussian one. Therefore $z(t)$ will asymptotically tend to vanish,

$$\lim_{t\to\infty} z(t) = 0$$

Actually we can go one step further and take advantage of the explicit form of the Gaussian in order to define the characteristic time in which the system will decohere. This is given as

$$\tau_D = \frac{1}{4g^2\sqrt{2N}|\alpha||\beta|} \tag{7.22}$$

Remarks

- As one could have probably guessed intuitively the decoherence time is shorter when the subsystems are strongly coupled and when the number of degrees of freedom of the environment - a measure of which is N - are increased.

- One could argue that the example under consideration is special, forbidding us to draw general conclusions. However this is not the case. Up to a change of basis in the tensor factors, the only special property of our Hamiltonian (although a property of great importance - see section 3.2) is the fact that we have chosen it to be of the form

  $$H_{se} = \sum_k H_1^{(k)} \otimes H_k \otimes I_{k'}$$

  where all $H_1^{(k)}$'s commute with each other

- One final remark is in order. There is a subtlety hidden in one of our approximations. Thecoefficient function $z(t)$ will not exactly die off after time τ_D. Actually it has some quasi-periodic properties. This traces back to equation (18), in which the periodic nature of $z(t)$ is apparent. Up to that point we had made no approximations. However in order to simplify things at that point we took the large N limit and we approximated the binomial distribution by a Gaussian. The end result is that, if we want to be careful, we should point out that the system will eventually recohere.

  The next logical question one would ask is whether this poses any threat to the explanation of quantum to classical transition through decoherence. The answer is no. Actually there is a similar effect in classical statistical physics, which we?ll mention very briefly:

  There is an argument against irreversibility of classical systems that takes its power from Poincare's reccurence theorem. The spirit of the theorem is that a finite system that starts evolving in time will return to its initial state infinitely many times. So, there is the cause of our problem. We assumed our system to be infinite, although realistic systems in physics are always finite.

  The counter-argument to the spoiling of decoherence due to recoherence is that typical Poincare times are of the order of $N!$. If we stop for a moment now to consider all the facts we end up with two important considerations: For a physical system, for which N is typically of the order 10^{23}, on the one hand there is extremely rapid decoherence, as commented above, and on the other hand extremely slow recoherence (just to immediately decohere again). The conclusion is that for all practical purposes the system decoheres and becomes classical.

7.3. Pointer Basis

The pointer basis is defined as the basis in which the RDM is diagonal. As we will see, this basis is selected by the environment.

7.3.1. Quantum Measurement revised

Suppose we want to perform a measurement on a quantum state $|\psi\rangle$. And assume that initially there are no correlations between the state vector of the system and that of our measurement apparatus.

$$|A_0\rangle \otimes |\psi\rangle = \sum_s a_s |A_s\rangle \otimes \sum_l |l\rangle \tag{7.23}$$

Now, likewise equation (9), according to the von Neumann pre-measurement scheme, the unitary time evolution suffices to establish correlations between the two.

$$|A_0\rangle \otimes |\psi\rangle \to \sum_s |A_s\rangle \otimes |s\rangle \tag{7.24}$$

One can argue that the observable that is measured is

$$\hat{S} = \sum_s e_s |s\rangle \langle s|$$

However, we can express the state of the apparatus in an alternative orthonormal basis. For instance,

$$|A_r\rangle = \sum_s \langle A_s | A_r\rangle |A_s\rangle \tag{7.25}$$

In this case, the final state would be,

$$\sum_s c_s |A_s\rangle \otimes |s\rangle = \sum_r |A_r\rangle \otimes \sum_s c_s \langle A_s | A_r\rangle |A_s\rangle = \sum_r d_r |A_r\rangle \otimes |r\rangle \tag{7.26}$$

with

$$d_r |r\rangle = \sum_s c_s \langle A_s | A_r\rangle |s\rangle$$

Now, one can say that the observable that is measured by the apparatus has the form,

$$\hat{R} = \sum_r f_r |r\rangle \langle r|$$

Does that imply that, after the measurement, the apparatus contains information about both observables? This is contradictory, because the two observables do not, in general, commute! The only way for our description to be consistent with the fact that real devices measure a specified physical quantity is that a decomposition of the form of equation (24) to be unique. The question that arises is, "What determines the unique pointer basis of the apparatus which records the corresponding states of the system?".

7.3.2. Environment-Induced-Superselection

Consider now the case where we take into account the interaction of the measurement apparatus with a third quantum system $|E_0\rangle$, which represents the environment. Then the total state will evolve as,

$$|E_0\rangle \otimes |A_0\rangle \otimes |\psi\rangle \underset{H_{ea}}{\rightarrow} \left[\sum_p b_p |E_p\rangle \otimes |A_p\rangle\right] \otimes |\psi\rangle \underset{H_{as}}{\rightarrow} \sum_p b_p |E_p\rangle \otimes |A_p\rangle \otimes |p\rangle \quad (7.27)$$

where H_{ea} and H_{as} represent the interaction hamiltonians between the environment with the apparatus and the apparatus with the system, respectively. At this point, a uniqueness theorem solves the ambiguity. The Tridecompositional Uniqueness Theorem states that if we can decompose the Hilbert space to a number of subspaces that is greater than or equal to three, as in equation 27, then its uniqueness is guaranteed. However it does not state anything about which is the basis in which the diagonalization takes place. And up to recent developments there is no reliable way to determine it formally.

The only way to say something about the proper basis is through physically reasonable arguments. For example, it would make sense to state that the pointer basis should be the one in which the system-apparatus correlations $|s_n\rangle |a_n\rangle$ are left undisturbed under the subsequent formation of correlations with the environment (the stability criterion). One can then find a sufficient criterion for dynamically stable pointer states that preserve the system-apparatus correlations in spite of the interaction of the apparatus with the environment by requiring all pointer state projection operators

$$\hat{P}_n^{(A)} = |a_n\rangle \langle a_n|$$

to commute with the apparatus-environment interaction Hamiltonian, i.e.,

$$[\hat{P}_n^{(A)}, \hat{H}_{ae}] = 0 \quad \text{for all } n \quad (7.28)$$

Other criteria similar to the commutativity requirement have been suggested for the selection of the preferred pointer basis, because it turns out that in most realistic cases the simple relation of equation (7.28) can only be fulfilled approximately. Alternative criteria that have been suggested, have been based for example on the von Neumann entropy or the purity $(\mathrm{Tr}\rho_\psi^2(t))$.

In order to attain a physical feeling of the spirit of the aforementioned arguments let us proceed to the following example:

Consider conducting some quantum experiment, for example an experiment involving electron measurement or bubble chambers etc. These sorts of experiments have several natural enemies. In particular, air molecules or random photons (sourced for example from cosmic microwave background radiation) could scatter o? the electrons that are relevant with the experiment.

Let us now ask ourselves: What does this result into?

Restricting to radiation it is sensible to state that by scattering with the electrons under study, the latter are practically being constantly monitored as long as their position is concerned. In this sense the environment itself imposes on the system that the preferred pointer basis is in the position space. This property holds in general. Depend- ing on the specific system under study and its interaction with the environment, other quantities could be monitored. One can find lots of examples to convince him(her)self regarding the validity of this statement. This mechanism is known as environment-induced superselection and it describes (as already mentioned) how the environment determines the pointer basis.

8. Consistent Histories - Hao Wu

8.1. Introduction

A theory is considered to be something that can answer certain kind of questions. Questions are asked by observers. We assume that asking questions is not really a physical process, it doesn't mean any change in the universe. However, not every set of questions can have a reasonable answer, for example "which slit the photon passes through" in the double-slit experiment. So we need a tool to check if the set of questions we ask is compatible. Because of the non-deterministic nature of the quantum mechanics, the answer to the questions is only probabilistic. A compatible set of questions has a well-defined probabilistic answer while for an incompatible set of questions the probability can not be assigned reasonably to each possible outcomes.

Now we introduce a theory which provides us a criterium for the compat-

ible set of questions and tells us how to calculate the probabilities of each possible outcomes in that case. It?s called "Consistent Histories" and is an interpretation of quantum mechanics, which means that it gives exactly those predictions of physical phenomena as the standard quantum mechanics will give. In this sense, it?s not "another" theory, and all the different interpretations of quantum mechanics are equally true. The value of a certain interpretation therefore completely lies in its underlying philosophy which may inspire physicists to find a really new theory, and the potential of its formulation to be extended to a really new theory. In the case of classical mechanics we have newtonian, lagrangian, hamiltonian and hamilton-jacobian formulations. They are equally true in classical mechanics, but they are not the same as we know.

This chapter is based on [13] and [11].

8.2. History Hilbert Space and Sample Space

We describe the system which we are interested in by a Hilbert space. Since we want to know the property of a system not only at a single time but in a sequence of time points, we need a history Hilbert space.

Definition 1 *For a given sequence of time points* $(t_i)_{i=1,\dots,f}$, $\breve{H} = H_1 \odot H_2 \odot \cdots \odot H_f$ *is called a history Hilbert space, where* H_i *is the Hilbert space of the system for all* i. The symbol $\odot$ means the direct product in time order, and the symbol $\otimes$ is used to denote the direct product in the same time.

Our question is actually a set of "yes/no"-type questions. For example if we want to know the position, we can divide the whole space into left part and right part and then ask "is the position in the left part?" and so on. For doing this we introduce the history and the sample space.

Definition 2 *A projector in the history Hilbert space*

$$Y = F_1 \odot F_2 \odot \cdots \odot F_f$$

is a history where F_i *is a projector in* H_i.

We denote the projector onto the "spin-up" state of a spin-1/2 particle with $[z^+]$ and so on. Then $[z^+]\odot[x^-]\odot[x^+]$ is an example of history. The physical

meaning of it is that the particle is "spin-up" at time t_1, "spin-left" at time t_2 and "spin-right" at time t_3.

Definition 3 *The set of histories constructed by decomposition of history identity*

$$\{Y^\alpha\} : \breve{I} = \sum_\alpha Y^\alpha$$

is a sample space, where the history identity $\breve{I}$ is the identity projector in the history Hilbert space.

Now we can formulate our questions by specifying the sample space and then asking about the probability of each history in it.

For example

$$\begin{aligned} Y^0 &= [z^-] \odot I \odot I \\ Y^1 &= [z^+] \odot [x^+] \odot [z^+] \\ Y^2 &= [z^+] \odot [x^+] \odot [z^-] \\ Y^3 &= [z^+] \odot [x^-] \odot [z^+] \\ Y^4 &= [z^+] \odot [x^-] \odot [z^-] \end{aligned} \tag{8.1}$$

is a 3-times sample space of a spin-1/2 particle. This sample space is usually to describe the system with initial state $[z^+]$. Y^0 simply has 0 probability but has to be included in the sample space for the completeness. If we assume the time evolution is just the trivial time evolution, i.e., the Hamiltonian is constantly say 0, this example actually describes the Stern-Gerlach version of double-slit experiment, where the second Stern-Gerlach magnet has been rotated by 90 degrees.

For Y^α and Y^β from the same sample space we have

"not Y^α" $\quad Y^{\tilde{\alpha}} = \breve{I} - Y^\alpha$

"Y^α and Y^β" $\quad Y^\alpha \wedge Y^\beta = Y^\alpha Y^\beta = 0$

"Y^α or Y^β" $\quad Y^\alpha \vee Y^\beta = Y^\alpha + Y^\beta$

Now look at Y^1 and Y^3 in example(1). They correspond to the two cases "particle passes through the left slit" and "particle passes through the right

slit". From quantum mechanics we know that these cases interfere with each other, so we may suppose that these two histories are not consistent with each other and the set of questions specified by this sample space is not compatible.

8.3. Born Rule

The question that we want to ask corresponds directly to the choice of sample space. Quantum mechanics can only offer a probabilistic answer which means classical probabilities have to be assigned to each history in the sample space. Though the histories in a sample space are mutually exclusive and sum up to history identity, it?s not guaranteed that classical probabilities can be assigned to them.

Consider a 2-times history of form

$$Y^\alpha = [\phi_0] \odot [\phi_1^\alpha]$$

and denote the probability weight of it by $W(Y^\alpha)$. Assume the system is isolated. The Born rule is

$$\begin{aligned} W(Y^\alpha) &= |\langle \phi_1^\alpha | T(t_0, t_1) | \phi_0 \rangle|^2 \\ &= \cdots \\ &= \mathrm{Tr}([\phi_0] T(t_0, t_1) [\phi_1^\alpha] [\phi_1^\alpha] T(t_0, t_1) [\phi_0]) \end{aligned} \tag{8.2}$$

where $T(t_0, t_1)$ is the unitary time evolution of the isolated system.

We see that it is important to know the time evolution T, or say the Hamiltonian of the system for calculating the probabilities.

We denote $[\phi_1^\alpha] T(t_0, t_1) [\phi_0]$ by the chain operator $K(Y^\alpha)$ and $\mathrm{Tr}(A^\dagger B)$ by the scalar product $\langle A, B \rangle$, then

$$W(Y^\alpha) = \mathrm{Tr}(K^\dagger(Y^\alpha) K(Y^\alpha)) = \langle K(Y^\alpha), K(Y^\alpha) \rangle \tag{8.3}$$

For a general history of form

$$Y = F_1 \odot F_2 \odot \cdots \odot F_f \tag{8.4}$$

we define the chain operator as following.

Definition 4 *The chain operator of the history Y of the form (4) is*

$$K(Y) = F_f T(t_f, t_{f-1}) F_{f-1} \cdots T(t_1, t_0) F_0$$

and

$$K(Y + Y') = K(Y) + K(Y')$$

Then for a general history Y^α, equation (3) gives the probability weight to it. This is the general form of the Born rule.

8.4. Consistency Condition

As mentioned in the previous section, classical probabilities can not always be assigned to the histories in a sample space by using the Born rule. Classical probability obeys the additivity, which means

$$W(E \wedge F) = W(E) + W(F)$$

$\forall$ different E, F in sample space.

Now we take a sample space $\{Y^\alpha\}$ and consider a projector Y of the form

$$Y = \sum_\alpha \pi^\alpha Y^\alpha$$

where $\pi^\alpha = 0, 1$.

$$\begin{aligned} W(Y) &= \left\langle \sum_\alpha \pi^\alpha K(Y^\alpha), \sum_\beta \pi^\beta K(Y^\beta) \right\rangle \\ &= \sum_\alpha \sum_\beta \pi^\alpha \pi^\beta \langle K(Y^\alpha), K(Y^\beta) \rangle \end{aligned} \tag{8.5}$$

If the additivity holds, we should have

$$W(Y) = \sum_\alpha \pi^\alpha \langle K(Y^\alpha), K(Y^\alpha) \rangle \tag{8.6}$$

To equal (6) and (7), we have to require

$$\langle K(Y^\alpha), K(Y^\beta) \rangle = 0 \,, \ \forall \alpha \neq \beta \tag{8.7}$$

Equation (8) is the consistency condition for the sample space $\{Y^\alpha\}$.

Sample spaces, for which the consistency condition is satisfied , are called consistent sample spaces. Classical probabilities can only be used for consistent sample spaces. If sample space is not consistent, one can observe quantum superposition(interference) of histories.

We have to notice that the consistency condition can only be applied on an isolated system since it's required in the definition of the chain operator, in which the unitary time evolution operator can not be defined in an open system. So there are two ways to construct a consistent sample space. The first strategy is to ask questions smartly: choose a smart way(you should not ask about everything) to decompose the history identity. The second strategy is to bring something(measurement apparatus, environment..) into interaction(in particular entanglement) with the system so that we have to enlarge our Hilbert space. Then we can ignore a part of the whole system. The time evolution T is changed in this case.

8.5. Examples

A very important example is shown in (1). We assume that spin-1/2 particle is isolated, which means that it is not interacting with anything. Then the unitary time evolution operator is just trivially 1. A simple calculation shows

$$\langle K(Y^1), K(Y^3)\rangle = \frac{1}{4} \tag{8.8}$$

and

$$\begin{aligned} 1 &= W(Y^1 + Y^3) \\ &= W(Y^1) + W(Y^3) + 2\langle K(Y^1), K(Y^3)\rangle \\ &= \frac{1}{4} + \frac{1}{4} + 2 \times \frac{1}{4} \end{aligned}$$

We see that the sample space (1) is not consistent under the trivial time evolution and there is an interference between histories. This is just a Stern-Gerlach version of the double-slit experiment as mentioned before.

Now we measure the the spin of the particle in x-direction at time t_2. Then the particle is in interaction with the apparatus, so we have to enlarge the Hilbert space and the Hilbert space describing the apparatus has to be cut in. We assume that the measurement is fully efficient which means that we

can just use a two dimensional subspace of the huge Hilbert space describing the macroscopic apparatus and if we set our apparatus at time t_1 to the state $|X^+\rangle$, then we can specify the time evolution:

$$T(t_2,t_1) = \begin{array}{c} \\ x^+X^+ \\ x^+X^- \\ x^-X^+ \\ x^-X^- \end{array} \begin{array}{c} \begin{array}{cccc} x^+X^+ & x^+X^- & x^-X^+ & x^-X^- \end{array} \\ \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & & 1 \\ & & 1 & \end{pmatrix} \end{array}$$

$$T(t_3,t_2) = \begin{array}{c} \\ x^+X^+ \\ x^+X^- \\ x^-X^+ \\ x^-X^- \end{array} \begin{array}{c} \begin{array}{cccc} x^+X^+ & x^+X^- & x^-X^+ & x^-X^- \end{array} \\ \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \end{array} \tag{8.9}$$

But by the decomposition we don't want to ask questions about the apparatus(actually we cannot ask arbitrary questions, in this example we can choose the "X-basis" to decompose the identity for the apparatus but not the "Z-basis" as shown in the next example), so the identity projector onto the subspace of the apparatus will not be de- composed. We have our sample space:

$$\begin{aligned} Y^0 &= ([z^-]\otimes I + [z^+]\otimes[X^-])\odot(I\otimes I)\odot(I\otimes I) \\ A &= ([z^+]\otimes[X^+])\odot([x^+]\otimes I)\odot([z^+]\otimes I) \\ B &= ([z^+]\otimes[X^+])\odot([x^+]\otimes I)\odot([z^-]\otimes I) \\ C &= ([z^+]\otimes[X^+])\odot([x^-]\otimes I)\odot([z^+]\otimes I) \\ D &= ([z^+]\otimes[X^+])\odot([x^-]\otimes I)\odot([z^-]\otimes I) \end{aligned} \tag{8.10}$$

We check the consistency condition between A and C:

$$\begin{aligned} K^\dagger(A) &= ([z^+]\otimes[X^+])\cdot T(t_2,t_1)\cdot([x^+]\otimes I)\cdot 1\cdot([z^+]\otimes I) \\ &= \frac{1}{4}\begin{pmatrix}1&1\\1&1\end{pmatrix}\otimes\begin{pmatrix}1&0\\0&0\end{pmatrix}\cdot\begin{pmatrix}1&&&\\&1&&\\&&&1\\&&1&\end{pmatrix}\cdot\begin{pmatrix}1&0\\0&0\end{pmatrix}\otimes\begin{pmatrix}1&0\\0&1\end{pmatrix}\cdot\begin{pmatrix}1&1\\1&1\end{pmatrix}\otimes\begin{pmatrix}1&0\\0&1\end{pmatrix} \\ &= \frac{1}{4}\begin{pmatrix}1&0&1&0\\0&0&0&0\\1&0&1&0\\0&0&0&0\end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
K(C) &= ([z^+] \otimes I) \cdot 1 \cdot ([x^-] \otimes I) \cdot T(t_2, t_1) \cdot ([z^+] \otimes [X^+]) \\
&= \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & & 1 \\ & & 1 & \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\
&= \frac{1}{4} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&\Rightarrow \\
\langle K(A), K(C) \rangle &= \mathrm{Tr}(K^\dagger(A) K(C)) = 0 \\
W(A) &= \mathrm{Tr}(K^\dagger(A) K(A)) = \frac{1}{4} \\
W(C) &= \mathrm{Tr}(K^\dagger(C) K(C)) = \frac{1}{4}
\end{aligned} \tag{8.11}
$$

We see that A and C are consistent with respect to each other and both have probability weight $\frac{1}{4}$.

It can be shown by similar calculation that the sample space (11) is consistent under the time evolution (10) and each of the histories A, B, C and D has probability weight $\frac{1}{4}$.

Now we replace the apparatus in the previous example by a second spin-1/2 particle. The initial state of the particle is again in $|X^+\rangle$. The time evolution is the same as (10). At time t_2, if we still don't ask about the state of the second particle, then it's exactly the case as in previous. If we now ask about the spin of the second particle in Z-direction at time t_2 (consider an EPR-type experiment where Alice wants to know the spin of the first particle in x-direction and Bob wants to know the spin of the second one in Z-direction, but they have not done any measurement yet), then we decompose each of

the histories A, B, C and D in (11) into two histories:

$$Y^0 = ([z^-] \otimes I + [z^+] \otimes [X^-]) \odot (I \otimes I) \odot (I \otimes I)$$

$$\begin{aligned}
A^1 &= ([z^+] \otimes [X^+]) \odot ([x^+] \otimes [Z^+]) \odot ([z^+] \otimes I) \\
A^2 &= ([z^+] \otimes [X^+]) \odot ([x^+] \otimes [Z^-]) \odot ([z^+] \otimes I)
\end{aligned}$$

$$\begin{aligned}
B^1 &= ([z^+] \otimes [X^+]) \odot ([x^+] \otimes [Z^+]) \odot ([z^-] \otimes I) \\
B^2 &= ([z^+] \otimes [X^+]) \odot ([x^+] \otimes [Z^-]) \odot ([z^-] \otimes I)
\end{aligned}$$

$$\begin{aligned}
C^1 &= ([z^+] \otimes [X^+]) \odot ([x^+] \otimes [Z^+]) \odot ([z^+] \otimes I) \\
C^2 &= ([z^+] \otimes [X^+]) \odot ([x^+] \otimes [Z^-]) \odot ([z^+] \otimes I)
\end{aligned}$$

$$\begin{aligned}
D^1 &= ([z^+] \otimes [X^+]) \odot ([x^+] \otimes [Z^+]) \odot ([z^-] \otimes I) \\
D_2 &= ([z^+] \otimes [X^+]) \odot ([x^+] \otimes [Z^-]) \odot ([z^-] \otimes I)
\end{aligned} \tag{8.12}$$

We can use a similar calculation using the matrix notation to show

$$\begin{aligned}
\langle K(A^1), K(C^1) \rangle &= \frac{1}{8} \\
\langle K(A^2), K(C^2) \rangle &= -\frac{1}{8}
\end{aligned} \tag{8.13}$$

So the sample space (14) is not consistent.

We can also show:

$$\begin{aligned}
\langle K(A^1), K(C^2) \rangle &= 0 \\
\langle K(A^2), K(C^1) \rangle &= 0
\end{aligned}$$

Combined with (15) and linearity of $K(\cdot)$ and $\mathrm{Tr}(\cdot)$, we just get another proof of (13).

Now we see that if we ignore the questions about the spin of the second particle in Z-direction, we have a consistent sample space (11), if not, we get an inconsistent sample space (14). A classical probability distribution can not be assigned to (14), so our questions have no reasonable probabilistic

answer.

There is another way to get a consistent sample space from (14): we ignore the questions about the spin of the first particle in x-direction instead of questions about the spin of the second particle in Z-direction. So we sum up A^1 and C^1, A^2 and C^2, B^1 and D^1, B^2 and D^2.

$$\begin{aligned} Y^0 &= ([z^-] \otimes I + [z^+] \otimes [X^-]) \odot (I \otimes I) \odot (I \otimes I) \\ Y^1 &= ([z^+] \otimes [X^+]) \odot (I \otimes [Z^+]) \odot ([z^+] \otimes I) \\ Y^2 &= ([z^+] \otimes [X^+]) \odot (I \otimes [Z^+]) \odot ([z^-] \otimes I) \\ Y^3 &= \otimes [X^+]) \odot (I \otimes [Z^-]) \odot ([z^+] \otimes I) \\ Y^4 &= \otimes [X^+]) \odot (I \otimes [Z^-]) \odot ([z^-] \otimes I) \end{aligned} \tag{8.14}$$

A similar calculation can show that the sample space (16) is consistent. The probability weight of them are:

$$\begin{aligned} W(Y^0) = W(Y^2) = W(Y^3) = 0 \\ W(Y^1) = W(Y^4) = \frac{1}{2} \end{aligned}$$

We see that to get a consistent sample space, we either ignore the questions about the x-spin of first particle or the questions about the Z-spin of the second particle. If we ask both together, then our sample space is not consistent which means there is no reasonable answer to our questions. The "realities" of the two particles are dependent of the choice of our set of questions. But remember that asking questions is not a real physical process. So the "Consistent Histories" is an explicitly non-realistic interpretation of quantum mechanics, where we use the term "realistic" as in the discussion about EPR-paradox.

9. Many worlds - Maximilian Jeblick

9.1. Introduction

One of the most stunning and daring proposal in the history of modern physics was made by Hugh Everett in 1957, when he formulated the Many Worlds (MW) Interpretation of quantum mechanics. Despite the fact that the phrase "many worlds" was only introduced about 10 years later by Bryce

de Witt, Everett nevertheless can be considered as the founding father of MW. For more historical details and funny anecdotes I strongly recommend the BBC documentary "Parallel worlds, parallel lives?" (→ Youtube).

Despite its thrilling consequences, which will be discussed later, MW is a simple con- sequence of standard quantum mechanics plus the assumption that the collapse of the wave function is unphysical and in contradiction with unitary time evolution. Indeed the "collapse" of the wave function during a "measurement"[1] is highly unsatisfactory, contradicts with the Schrödinger Equation (SEQ) and is moreover ad hoc. Assuming that the whole universe is guided by quantum mechanics, von Neumann pointed out that the linearity of the Schrödinger Equation does not allow any collapse. Thus it is reasonable to ask if it possible to modify Copenhagen Quantum Mechanics in such a way, that there is no need for any collapse of the wave function. The astonishing answer Everett gives is: Yes! It is possible to withdraw the collapse postulate without adding further entities!

9.2. Sketch of the Many World Interpretation

As we figured out we will not modify the usual quantum formalism despite the fact that we assume the SEQ to be valid for all times. In order to formulate a measurement we will apply the von Neumann Scheme. As the most simple example we will consider the spin measurement of a spin $1/2$ particle. The initial state reads as (omitting the spatial part of the electron wave function):

$$|\Psi(t_i)\rangle = (\alpha\,|\uparrow\rangle + \beta\,|\downarrow\rangle) \otimes \left|\Psi^A[0]\right\rangle \tag{9.1}$$

where $|\Psi^A[0]\rangle$ denotes the apparatus in its ready state "0". After a while the electron will hit the apparatus and enforce an interaction between the apparatus and the electron[2]. Hence the final configuration will read as follows:

$$|\Psi(t_f)\rangle = \alpha\,|\uparrow\rangle \otimes \left|\Psi^A[\uparrow]\right\rangle + \beta\,|\downarrow\rangle \otimes \left|\Psi^A[\downarrow]\right\rangle \tag{9.2}$$

This result, which is of course well-known, does not seem to be satisfactory at all! What does the apparatus show? We have arrived at Schr??inger?s cat paradox. However there exists a loophole: If we consider the summands in (2) as distinct entities, we might also say: One outcome is realized in one

[1] For are detailed criticism about "measurement" see also [4]

[2] For more details cf. any book with a detailed description of the von Neumann Scheme

branch and the other outcome is realized in another branch. If we consider the apparatus coupled to a human being (that is we read off the apparatus pointer), we might construct the following sequence:

$$|\Psi(t_f)\rangle = \alpha\,|\uparrow\rangle \otimes \left|\Psi^{A}[\uparrow]\right\rangle \otimes \left|\Psi^{Human}[\uparrow]\right\rangle + \beta\,|\downarrow\rangle \otimes \left|\Psi^{A}[\downarrow]\right\rangle \otimes \left|\Psi^{Human}[\downarrow]\right\rangle \quad (9.3)$$

That is within one branch apparatus and human agree upon the measurement result. Moreover any interaction which *might* be used as an honest measurement apparatus enforces effective decoherence [8] . According to MW there exist different branches and **within** each branch everything is self-consistent. The branches don't "know" from each other, i.e., decoherence forbids any further interaction or "communication" between different branches[3]. In this sense decoherence generates an effective collapse which is FAPP indistinguishable from a real existing collapse[4].

A short remark to the name "Many Worlds":

As seen from (2) or (3) we have indeed a many worlds interpretation, that is there exist different distinct entities where each might be considered as a distinct "universe". Indeed we might label "system spin up" as universe 1, where one reads "spin up" and writes down "spin up" and universe 2, where one writes down "spin down". In my opinion the phrase "many worlds" is not only a pictorial metaphor, since we have distinct, FAPP non interacting, physical systems and each of these systems consists of entities which strongly imply the word "universe"[5].

9.3. Quantum mechanics as a partly interpreted system

We will now present a more general procedure, which allows to analyze QM on a more fundamental level. While this analysis is not specific to MW it

[3]Everett therefore used the phrase "relative state formalism" instead of many worlds. Since the decomposition into different branches is not unique one needs to investigate this problem on a more profound level. See also the chapter on decoherence.

[4]There are attempts to design experiments which could distinguish between an effective and a dynamical (GRW) collapse.

[5]This resembles a collection of "classical" universes with different realizations of experiments. One also might consider the universal wave function as "the universe" paying the price that everything we have experienced so far is only realized within one branch. That is, in this reading it is a priori impossible to penetrate into regions of the universe which are realized in other branches.

nevertheless helps to distinguish more systematically different formulations of QM. We assume that a physical theory (as QM) is a partly interpreted formal system. That is, we split our theory as follows: The underlying axiomatic system is understood purely mathematical, denoted by F (formalism). The formalism solely consists of mathematical axioms and theorems (see MQM!). We define rules of correspondence[6], which identify objects in F with physical reality[7]. As an example, the parameter "t" in $\Psi(x,t)$ is supposed to represent time, i.e., the thing we read off from a clock in an experiment. A physical theory T is therefore to be understood as consisting of two parts, namely $T = F_R$. F_R may be thought as a machine, which should represent as many physical experiments as possible. One the one hand we calculate stuff using the formalism F and compare the resulting numbers with numbers we get from the experiment using the rules of correspondence R. Several philosophers claim that physics must not consist of more than that: Comparison of experimental numbers with theoretical numbers. Yet we do NOT have any interpretation, i.e. yet we must NOT talk about matter deflecting particles, etc. An interpretation, denoted by M (model), is understood to provide a "picture" or understanding of F_R. M is supposed to be isomorphic to F_R, that is contains (at least) all elements of F_R[8]. If one watches any physical documentary on TV one barely sees any formulas explained (the formalism F), but computer generated animations. This is the model one constructs. Thus different interpretations of QM might either differ F, in R or in M.

If we consider the usual (=Copenhagen) postulates of QM, it is very reasonable to criticize the collapse postulate due to inconsistencies. MW can be constructed using two steps:

In order to abandon the collapse, we assume the new formalism $F_R(QM/C)$, where $F_R(QM/C)$ = usual QM without collapse. This seems like a reasonable step. As a second step it is claimed that the resulting model M **HAS** to be a MW interpretation, i.e., the formalism enforces its own interpretation[9].

[6]These are expressed in so-called meta language.

[7]Some parts of F might not correspond with physical reality, therefore the theory may only be partly interpreted.

[8]If M consists of more elements, then this might be an indication for altering F_R in order to achieve a stronger theory T'.

[9]Since Born's law is claimed to follow as a theorem we can omit this axiom, denoting the modified theory again $F_R(QM/C)$

The argumentation goes as follows:

We know that usual QM is tested with high precision. However the collapse postulate is inconsistent with quantum cosmology, since there is no external observer. It is assumed that the idea of quantum cosmology is reasonable. Therefore we eliminate this postulate but do not add anything else (Occam's razor). Applying the von Neumann Scheme we observe that EVERY possible outcome of an experiment is realized. Hence we identify each branch of the universal wave function with a different world.

9.4. Properties of the Many World Interpretation

Well, there is not much to say, except about Born's law. As we have seen, we do not alter the formalism in a radical way. Hence all results from Copenhagen QM can be adapted. Decoherence generates an effective collapse, which replaces the collapse postulate. As mentioned, the decomposition into different branches, which might not be unique, can be handled using either some classical limiting procedure or environment-induced superselection. In most cases this problem seems to be treatable unambiguous. However the striking question which remains is: How is it possible that deterministic MW can reproduce Born's law? Or more drastically: Is it even *possible* to formulate a consistent MW theory which is compatible with Born's law? The author believes that this is not possible, thus dismissing MW as an honest alternative interpretation.

9.5. Statistics

In this section one early attempt to derive Born's law is presented. It has been claimed that MW is able to recover the usual statistic without assuming further propositions. Usually one considers an ensemble, consisting of N identical subsystems coupled to a measurement apparatus. For simplicity we consider a spin $1/2$ particle. The initial (prepared) state reads as follows:

$$\begin{aligned} |\Psi_N(t_i)\rangle =& (\alpha\,|\uparrow_1\rangle + \beta\,|\downarrow_1\rangle) \otimes (\alpha\,|\uparrow_2\rangle + \beta\,|\downarrow_2\rangle) \otimes .... \\ & \otimes (\alpha\,|\uparrow_N\rangle + \beta\,|\downarrow_N\rangle) \otimes \left|\Psi^A[0,0,...,0]\right\rangle \end{aligned} \tag{9.4}$$

The state vector $|\Psi^A[0,0,...,0]\rangle$ is understood to act as a measurement apparatus (von Neumann Scheme) with N free "slots". The dynamics of such a

measurement is determined by applying the suitable unitary time evolution operator, i.e.,

$$|\Psi_N(t_f)\rangle = \hat{U}(t_f - t_i)\,|\Psi_N(t_i)\rangle \tag{9.5}$$

As a result we obtain:

$$\begin{aligned}|\Psi_N(t_f)\rangle = \alpha^N\,&|\uparrow_1\rangle \otimes |\uparrow_2\rangle \otimes |\uparrow_3\rangle \otimes .... \otimes |\uparrow_N\rangle \otimes \left|\Psi^A[\uparrow_1, \uparrow_2, \uparrow_3, ..., \uparrow_N]\right\rangle \\ + \alpha^{N-1}\beta\,&|\uparrow_1\rangle \otimes |\downarrow_2\rangle \otimes |\uparrow_3\rangle \otimes .... \otimes |\uparrow_N\rangle \otimes \left|\Psi^A[\uparrow_1, \downarrow_2, \uparrow_3, ..., \uparrow_N]\right\rangle \\ + \alpha^{N-1}\beta\,&|\uparrow_1\rangle \otimes |\uparrow_2\rangle \otimes |\downarrow_3\rangle \otimes .... \otimes |\uparrow_N\rangle \otimes \left|\Psi^A[\uparrow_1, \uparrow_2, \downarrow_3, ..., \uparrow_N]\right\rangle \\ + ... + \beta^N\,&|\downarrow_1\rangle \otimes |\downarrow_2\rangle \otimes |\downarrow_3\rangle \otimes .... \otimes |\downarrow_N\rangle \otimes \left|\Psi^A[\downarrow_1, \downarrow_2, \downarrow_3, ..., \downarrow_N]\right\rangle\end{aligned} \tag{9.6}$$

These are in total 2^N terms. According to MW each term represents a different universe. Notice that the state $|\Psi_N(t_f)\rangle$ is L^2-normalized to 1. This can be both seen from the unitary time evolution as well from an explicit calculation:

$$\langle \Psi_N(t_f)\,|\,\Psi_N(t_f)\rangle = \sum_{k=0}^{N} \binom{N}{k} |\alpha|^{2(N-k)} |\beta|^{2k} = (|\alpha|^2 + |\beta|^2)^N = 1 \quad \forall N$$

We now regroup these "worlds" according to their relative frequencies for obtaining n-times spin up. Using this regrouping we want to show as $N \to \infty$ that the state $|\Psi_N(t_f)\rangle$ approaches the normalized vector $|\Psi(n = N|\alpha|^2)\rangle$, which represents all universes with the right quantum statistics. The regrouping reads as follows:

$$|\Psi_N(t_f)\rangle = \sum_{n=0}^{N} c_n\,|\Psi(n)\rangle \tag{9.7}$$

where

$$\begin{aligned}|\Psi(n)\rangle = \frac{1}{\Omega} \sum_{permutations\,\sigma} &|\uparrow_{\sigma_1}\rangle \otimes |\uparrow_{\sigma_2}\rangle \otimes ..... \otimes |\uparrow_{\sigma_n}\rangle \otimes |\uparrow_{\sigma_{n+1}}\rangle \\ &\otimes \left|\Psi^A[\uparrow_{\sigma_1}, \uparrow_{\sigma_2}, ..., \downarrow_{\sigma_N}\right\rangle\end{aligned} \tag{9.8}$$

(While mathematically not quite correct, this expression is hopefully self-explanatory.) The normalization $\Omega = \sqrt{\binom{N}{n}}$ is such that $\langle \Psi(m)\,|\,\Psi(n)\rangle = \delta_{mn}$. Therefore the coefficient c_n is equal to $\sqrt{\binom{N}{n}}\alpha^n\beta^(N-n)$. It remains to show that $|\Psi_N(t_f)\rangle$ approaches $|\Psi(n = N|\alpha|^2)\rangle$ for large N in L^2-norm. This is

equivalent in showing that $|c_n|^2$ approaches 1 if $n = N|\alpha|^2$. The underlying idea now seems to become more obvious: For all spins to be up, there exists only one universe with prefactor α^N which goes to 0, as $N \to \infty$ (since $\alpha < 1$ for nontrivial discussion). For having all except one spin up, one part of the prefactor reads as $\alpha^{N-1}\beta$ which also goes to zero . However the number of universes in which this statistic shows up is $\binom{N}{1} = N$, which goes to infinity but certainly not fast enough for compensation. In general n,the number of spins up, also grows when N grows rather than being fixed. Thus in general n is a function of N (In our proof we don't assume this (reasonable) property, since it will come out naturally). Balanced right the binomial (=number of universes) will compensate the zero sequence consisting of α's and β's. The universes which "add up" in just the right way are exactly the universes, which reproduce Born's law! Let?s state our idea a bit different: We have the time-evolved state, which has unity norm for all N. We decompose this vector into "relative frequencies"-states, which are also normalized to one for all N. For N going to infinity the time-evolved state will be identical to the "relative frequency"-state, which shows the right statistic.

So what does $|c_n|^2$ look like?

$$|c_n|^2 = \binom{N}{n}|\alpha|^{2n}|\beta|^{2(N-n)} = \binom{N}{n}(|\alpha|^2)^n(1-|\alpha|^2)^{(N-n)}$$

This is the binomial distribution, where $|\alpha|^2 = p$, i.e., the probability for getting $\frac{n}{N}$-times spin up. Hence, in the limit $N \to \infty$, the terms $c_n|\Psi(n/N)\rangle$ which do not show the right statistic will die out. This follows directly from the properties of the binomial distribution.

Remark: The whole proof can be made mathematical rigorous (working with the usual ϵ-criterion). Also, more general situations (not just spin 1/2) can be analyzed in a similar way.

Thus in a "typical universe" an infinite ensemble shows up to have the right statistic. There are at least to remarks, which need to be made:

1. A naive counting of worlds would NOT recover Born's law. If we naively count the number of universes, we would conclude that in most universes 1/2 of all events is spin up and 1/2 is spin down. In performing this counting however one implicitly assumes that the probability for

an observer to be in one of the two universes after a single measurement is equally distributed, i.e., also $1/2$.

2. Unfortunately this "proof" was dismissed even by the author (Bryce de Witt). Implicitly one assumes, by normalizing $|\Psi(n)\rangle$ in L^2-sense, Born's law. Moreover it is unclear how one can consistently transfer this "measure of typicality" to a finite N-ensemble. Other proposals to prove Born's law were made, but they have some severe flaws. It is even unclear how one can even enforce any statistics, if **ALL** possible outcomes are realized. How can one speak of probability if, when one flips a coin, he receives head in universe 1 AND tail in universe 2 at the same time? Not only severe philosophical problems arise (such as the problem of transtemporal identity), but also the fundamental absent of any mechanism, which selects a measurement result, is highly in contradiction with **ANY** statistics. We won't follow this deep discussion any further, for more information cf.....

9.6. Discussion of MW

We've seen that MW arises somehow "naturally" rom well known QM, if one abandons the collapse postulate. We will first discuss how "naturally" the implication $F_R(QM/C) \to M =$ MW is. In fact there exist other interpretations M', which also assume $F_R(QM/C)$. As an example the bare theory denies splitting by denying reality (in a very strong sense). From a "common sense" these theories are even more weird than MW, so we will not discuss them. Another attempt is to change $F_R(QM/C) \to F'_R(QM/C)$ not by adding any further postulates but to alter the very fundamental mathematical axiom of truth, namely Boolean algebra. This procedure is called Quantum logic. We do not know how one could accept this procedure as being convincing without abandon usual common sense.

Thus it seems - assuming $F_R(QM/C)$-MW interpretation is by far the most convincing compared to the theories stated above. We will now list some pro's and con's for this.

Pro's:

- Assuming $F_R(QM/C)$ is not unreasonable at the first place, especially the collapse can be derived rather than needed to be postulated .

- Assuming $F_R(QM/C)$, MW seems to be the most convincing theory when tested against reasonable reasoning (this is however open to discussion!). - F is a minimal set (Occam's razor applied on F).

Con's:

- The sheer immense number of universes seem highly counter-intuitive.

- The sheer immense number of universes seem to create entities beyond necessity for describing one (our) universe. (Occam's razor applied on M)

- Problem of preferred basis (see:Decoherence)

- The highly daring hypothesis of MW relies on assuming $F_R(QM/C)$. This however cannot be regarded as a very strong argument since there exist various $F'_R(QM/C)$ (e.g., Bohmian Mechnics), which do not require an infinite amount of universes.

- The most severe criticism is that MW might not be in agreement with experiment, since it seems to be mathematically inconsistent with any statistics.

Conclusion:

For claiming such a tremendous impact of our physical perception of reality, MW is definitely not based on THE striking argument, which would strongly demand this interpretation. Moreover, as can be seen above, MW claims to follow from Occam's razor thus being a inevitable consequence. This might be true on the side of the formalism F, but is not true on the side of the model M. As a physicists one has the duty to check when claiming any daring hypothesis, if he has strongly convincing arguments. This seems not to be the case.

One could argue that in order to resolve this problem (replace MW by some other theory) one has to include further axioms to F. This is indeed problematic, since F together with R is able to describe quantum mechanical experiments (ignoring the statistical debate) . F is minimalistic, however leads to a "non minimalistic" interpretation. If one wants to resolve MW, it seems one needs to alter F (rather than R or M given F(QM/C)). This might be: - nonlinear stochastic time evolution (GRW) - one graviton criterion (as

proposed by Penrose), or any other dynamical deterministic process which causes the collapse (to be developed) - add further parameters (hidden variables) which determine in which branch the actual universe "splits". In view of MW, Bohmian mechanics can be considered to provide such a mechanism.

10. Bohmian mechanics - Henry Hanson and Franz Thoma

10.1. Realistic Quantum Theories

When we view interpretations of quantum mechanics, we find that most physicists are happy with some interpretation that does not assume realism. That is to assume that the results of experiments have a one-to-one correspondence with the state of a physical object that exists in our universe. Theories that do not assume realism assume in a certain sense that our universe is bigger than what we measure, i.e., he parts of the wave function that don't find a reality in our perception, find a reality in the perception of an observer who is in the other part of the wave function. What these theories don't explain however is by what mechanism is chosen which branch of the wave function an observer finds itself in. More specifically the mechanism is not part of the theory. It is because of this, that some physicists consider theories assuming reality. One of those theories is Ghirardi-Rimini-Weber (GRW) theory and the other is Bohmian mechanics. We will discuss the latter here.

Bohmian mechanics assumes the existence of the (arguably) simplest physical object to represent observations, i.e., a point-like object called the particle. It is assumed that this object has only one property, namely position, and that there are a finite (or at least countable) number of them and that they change position according to some differential equations (which will be derived in the following sections). From the position of the particle it is possible to infer what the outcome of experiments will be, just as it would be possible in the non-realistic theories if we knew at all times which branch we were in. The concept of "particle" defines a mechanism which decides which branch of the wave function is physical, i.e., assuming the knowledge of the particle positions and the quantum mechanical state, every observer observes a uniquely determined branch of the wave function.

10.2. Derivation of the Guiding Equation

Bohmian Mechanics takes the concept of particles serious, in the sense that there is a differential equation that yields particle trajectories. We can derive this equation just from the Schrödinger equation,

$$i\partial_t\psi = \left(-\frac{\Delta}{2m} + V\right)\psi$$

The only thing we need, aside from particles, is the Quantum Equilibrium Hypothesis, which is nothing else than Born's statistical law $\rho = |\psi|^2$. This may seem trivial since this relation is well known and often used, but it plays a very different role in Bohmian mechanics, as we will see later. Until then, we take $\rho = |\psi|^2$ as a postulate (which is experimentally well justified). We start with the derivation of the continuity equation (note that for $\bar{\psi}$ we need the complex conjugated Schrödinger equation,

$$-i\partial_t\bar{\psi} = \left(-\frac{\Delta}{2m} + V\right)\bar{\psi}$$

$$\begin{aligned}
\partial_t\rho = \partial_t(\bar{\psi}\psi) &= \bar{\psi}\partial_t\psi + \psi\partial_t\bar{\psi} \\
&= -i\bar{\psi}\left(-\frac{\Delta}{2m} + V\right)\psi + i\psi\left(-\frac{\Delta}{2m} + V\right)\bar{\psi} \\
&= \frac{i}{2m}\nabla\cdot(\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi}) \\
&= -\frac{1}{m}\nabla\cdot\mathrm{Im}(\bar{\psi}\nabla\psi) =: -\nabla\cdot j^\psi
\end{aligned}$$

Thus we have determined the conserved probability current of the wave function ψ, j^ψ, with the continuity equation

$$\partial_t|\psi|^2 + \nabla\cdot j^\psi = 0 \tag{10.1}$$

This current lives on the same space as the wave function, i.e., on configuation space. Now, mathematically we can decompose this current as $j^\psi = \rho v^\psi$, where v^ψ is a vector field on configuration space.

Until here, we had exactly the same results as in ordinary quantum mechanics; but now we introduce particles that are described by their positions $Q_i(t)$ in physical space and straightforwardly interpret v^ψ as their velocity field,

$$\dot{Q}(t) = v^\psi = \frac{1}{m}\mathrm{IM}\frac{\bar{\psi}\nabla\psi}{\bar{\psi}\psi} \tag{10.2}$$

where $Q(t) = (Q_1(t), ..., Q_N(t))$ is a vector in configuration space.

This equation was already written down in 1927 by L. deBroglie at the Solvay conference. However, after sharp (but wrong) objections from W. Pauli, de-Broglie abandoned the theory, until in 1952 D. Bohm re-discovered it.

Thus we have in Bohmian mechanics two physical objects, the Schrödinger wave function $\psi(x,t)$ and the particle positions $Q(t)$ that obey the two differential equations

$$i\partial_t\psi = \left(-\frac{\Delta}{2m} + V\right)\psi \quad , \quad \dot{Q} = \frac{1}{m}\mathrm{Im}\frac{\nabla\psi}{\psi}$$

Unlike in classical mechanics, a Bohmian system is not fully determined by positions and momenta at a time t_0, but by positions and the wave function at t_0. Hence momentum does not play the same fundamental role in Bohmian mechanics as in classical mechanics; in fact, a Bohmian particle has no properties other than its position, while all other properties (mass, charge) and observables of the system belong to the wave function. In this sense, quantities like Energy and momentum (and also their conservation laws) lose their meaning at the level of particles.

Nota been: v^ψ is not unique, we can add an arbitrary curl term such that the continuity equation is satisfied (similar to a gauge transformation). Hence, Bohmian mechanics is an *example* (in fact, the simplest example) for an extension of quantum mechanics featuring deterministic particle trajectories, but is not the unique theory that tells us what is "really" going on on microscopic level.

On the Heisenberg Uncertainity Relation: The existence of an exact particle position does not violate Heisenberg's Uncertainity Relation $\Delta q \Delta p \geq \frac{1}{2}$: Δp is the "wave number width" of ψ and Δq its spatial extent; Δq has nothing to do with the actual particle position Q (this is why we strictly use the notation Q for particle positions and q for the space coordinate of the wave function).

Quantum Equilibrium Hypothesis: The continuity equation relies on the equilibrium condition $\rho = |\psi|^2$. But what happens in the case of quantum non-equilibrium? Then we had to scrap the continuity equation and hence our guiding equation. Of course, in a *typical* Bohmian universe, we typically

experience quantum equilibrium; So we have to justify the hypothesis: We live in a typical Bohmian universe. There are some papers on this issue[7], but for now we can simply accept it, since there is no evidence for quantum non-equilibrium.[1]

Non-locality: In his famous 1964 paper[3], J. S. Bell showed that any theory that gives the same statistical results as quantum mechanics must be either non-realistic or non-local. The experimental verification of this led to the conclusion that nature indeed is non-local (under the assumption that nature itself is realistic).[2] One important note is that this is not in contradiction with special relativity: There already exist toy models of Bohm-like theories that are fully Lorentz invariant without having to choose an explicit frame[19].

10.3. The Measurement Process in Bohmian mechanics

In Bohmian mechanics, the measurement problem is resolved in a straightfoward way. We describe both the system and and the apparatus Quantum-mechanically, i.e., by wave functions $\psi(x), \Phi(y)$ and particle configurations $X(t), Y(t)$, respectively. For the (macroscopic) apparatus, we can assume for all practical purposes that macroscopically distinct states Φ_i (pointer positions, dots on photographic plates, ...) have disjoint support, $\mathrm{supp}\Phi_i \cap \mathrm{supp}\Phi_j = \varnothing \ \forall i \neq j$.

Now, for a simple model assume a two-state system (ψ_1, ψ_2 with $\langle \psi_1 \,|\, \psi_2 \rangle = 0$) and a measurement apparatus with three (macroscopically distinct) states (e.g., pointer positions) Φ_0 (null position, nothing measured) and Φ_1, Φ_2. The apparatus should measure in which state the system is, i.e., we expect

$$\psi_i \Phi_0 \to \psi_i \Phi_i$$

for the Schrödinger time evolution. Now take a superposition state $\psi = c_1\psi_1 + c_2\psi_2$, $|c_1|^2 + |c_2|^2 = 1$; initially ($t = 0$), the whole system is in the state

[1] This is unlike in classical statistical physics: There we have plenty of evidence of non-equilibrium, so we do not live in a typical thermodynamical universe! On this issue, see [8] chapter 4 (typicality in Thermodynamics) and chapter 11 (in Bohmian mechanics).

[2] Bell's inequalities are a criterion on locality of correlations. Experiments have shown that nature indeed exhibits non-local correlations. Under the assumption that nature itself is realistic we can conclude the nature is non-local. See for example the afterword of [13], just to mention a "standard quantum mechanics textbook".

$\psi(x)\Phi_0(y)$, which is a product state, i.e., system and apparatus are not entangled. During the measurement process ($t \to T$), both become entangled. This is described by Schrödinger time evolution,

$$\psi(x)\Phi_0(y) \overset{t\to T}{\to} c_1\psi_1\Phi_1(y) + c_2\psi_2\Phi_2(y)$$

Now we look at the particle configurations $X(t), Y(t)$ at $t = T$. Since Φ_1 and Φ_2 have disjoint support, the cross terms vanish and we have

$$\rho(x) = |c_1\psi_1(x)\Phi_1(y)|^2 + |c_2\psi_2(x)\Phi_2(y)|^2$$

at $t = T$. Thus, if the pointer is in position 2, the configuration $Y(T)$ is in $\operatorname{supp}\Phi_2$, but then $(X(t), Y(T))$ is in $\operatorname{supp}\psi_2\Phi_2$, i.e., $X(T)$ must be guided by the *effective wave function* ψ_2 alone,

$$v^{\psi}(X(T)) = v^{\psi_2}(X(T))$$

The ψ_1 branch of the initial wave function on the other hand is "empty" (does not contribute to guiding the particle, *not effective*). Hence we we have for all practical purposes[3] exactly the same dynamics as if the system was in the state ψ_2 in the first place, i.e., we *effectively* describe the system by the "collapsed" wave function ψ_2. Thus in Bohmian mechanics the collapse of the wave function is not a physical process, but merely an act of convenience, neglecting branches of the wave function that are very unlike to become effective again for any timescale of interest. — We do not need a postulate of collapse. Alone the assumption that particles exist and have a position immediately yields a formalism that recreates the formalism of collapse without having to postulate it. —

[3]We can assume that ψ_1 is empty with the following restrictions:

(a) The overlap of the macroscopic wavefunctions is not necessarily exactly zero,but "only" negigibly small; hence we have a very small overlap term $2\mathrm{Re}(\bar{c}_1 c_2 \bar{\psi}_1\psi_2\bar{\Phi}_1\Phi_2)$ in the probability density ρ that allows for a pointer position $Y(T) \in \operatorname{supp}\Phi_2$ although $X(T)$ is still guided by ψ_1 (i.e., $v^{\psi}(X(T)) \neq v^{\psi_1}(X(T))$)

(b) By Poincarè recurrence, the neglected branch of the wave function can become effective again.

We can ignore this for all practical purposes since (a) the overlap is ridiculously small for $\sim 10^{26}$ random phases (decoherence!) and (b) the Poincarè recurrence time is ridiculously big compared to all relevant time scales.

10.3.1. So, what do we actually measure?

The basic idea in the above model was that what we see is the actual macroscopic configuration $Y(T)$ of the apparatus. Hence what the apparatus measures is the *actual* particle position $X(t)$ of the system; not necessarily the *exact* position, but "does the particle go up or down", "is there a photon in the detector or not" etc.

From the position we then can infer values for observables (coarse-graining functions), depending on the experimental settings[4]. For example, the "spin observable" can be obtained by measuring the position (i.e. where the particle hits the photographic plate) of a particle that has travelled through a Stern-Gerlach magnet (experimental setting). The observable itself has no intrinsic value, we need not only the system itself, but the entire experimental setup to determine its value.

Note that it is always the theory that determines what we can measure. When we pass to a purely statistical description, then we can of course talk about measuring an observable like $\langle\psi| A |\psi\rangle$. Compare to Newtonian mechanics, where we can only measure positions q and momenta p, as opposed to its statistical theory (thermodynamics), where we measure observables like temperature T or pressure P.

10.3.2. Can I prepare Q?

To prepare a system, we have to be able to set all initial conditions at a certain time. This is fairly easy in Newtonian mechanics: we can prepare all initial conditions in- dependently, and hence prepare the time evolution of any system by setting p and q accordingly. In Bohmian Mechanics this looks different:

We can naturally prepare the position of a particle by measurement. For example in a single slit experiment, any particle that passes through the slit has to be in a certain interval at a certain time. We can as well prepare the effective wave function of a system via controlled interaction: For example we can wait for a hydrogen atom to emit a photon, and infer from the photon energy the effective wave function of the electron. This looks very promising,

[4]The experimental settings are encoded in the Hamiltonian via constraints or external fields.

but unfortunately in Bohmian mechanics ψ and Q are *not* independent, so we cannot prepare arbitrary systems (ψ, Q): Whenever we prepare the position of a particle to a certain accuracy, we have to live with a wave function that destroys this accuracy after a short amount of time, and when we indirectly prepare a system in a certain effective wave function, we have to do this in a way that allows us to control interactions, which in turn limits the accuracy of position measurements.

So to give a short answer: We can prepare particle positions and effective wave functions in Bohmian mechanics, but not both independently at the same time.

10.4. Non-Newtonian behavior

Since Bohmian mechanics is a theory about particles, it may be tempting to try to imagine the dynamics of these particles in a classical Newtonian sense. In fact Bohmian trajectories are typically very different from their Newtonian counterpart. This is be- cause Newtonian particles are guided by the concept of forces, which means that particles move along straight lines in the absence thereof (Newtons 2nd law). Bohmian particles on the other hand are guided by the "guiding wave principle", which means that their dynamics are governed by the wave function. Only through the wave function do the particles feel the presence of a potential and more importantly they do not in general move along straight lines when there is no potential. I would like to discuss some of the non-Newtonian behavior of Bohmian particles with the help of some examples.

10.4.1. Hydrogen Atom

One of the simplest examples where we can see non-Newtonian behavior is the hydrogen atom. First we will take a look at the ground state of the hydrogen atom. The wave function goes like $\kappa e^{-r/c}$, where κ and c are real constants. Thus the wave function for the ground state is real. But the gradient of a real function is again real. Therefore the velocity field for this wave function is 0. This means that if the initial wave function is in the ground state and the initial position is x_0 then $X(t) \equiv x_0$, i.e., the particle doesn't move. The Newtonian picture for this situation would be solving the Newton equations for a central $\frac{1}{r}$ potential with the initial velocity $v_0 = 0$

(lowest energy state) and initial position x_0. This would lead to the particle falling into the nucleus. But because we know that hydrogen atoms in the ground state are stable, we can conclude that Newtonian mechanics gives us the wrong dynamics and that Bohmian mechanics gives us an unambiguous[5] alternative.

Secondly we will examine the first exited state. It is a straightforward, but lengthy calculation that shows that the Bohmian trajectories are circles around the nucleus[6]. The corresponding Newton analogon would be to have a $\frac{1}{r}$ potential with an angular momentum corresponding to the "angular momentum" of orthodox quantum mechanics. But a charged particle in Newtonian mechanics would radiate if it traveled in a circle making the trajectory unstable. This once again shows that it is possible to have particles without Newtonian behavior that still give an accurate description of nature.

As a final note I would like to add that for superposition states the trajectories are much wilder and therefore even less Newtonian (if one wants to talk about more or less Newtonian), but they give the right statistics for measurements and are continuous (and satisfy all the other properties the velocity field satisfies).

10.4.2. Stern-Gerlach Experiment

There are two issues I will tackle with this example. The first is to show that Bohmian trajectories are not in general straight lines when there is no potential and secondly I would like to show how Bohmian mechanics gets away with giving the particle only one property: It's position.

The wave function for one silver atom after it has passed through a Stern Gerlach magnet is derived by solving the Pauli equation for the potential given by a stern Gerlach magnet and then letting the wave function evolve freely. Just as in any other interpretation of quantum mechanics, the wave function is a spinor. What makes Bohmian mechanics different is that the "spin of a particle" is *not* an intrinsic property of the particle. Therefore the

[5]But not necessarily unique, since we can add a curl term to the velocity field without changing the statistics.

[6]Of course, all eigenfunctions of a real Hamiltonian can be chosen real, yielding $v^\psi = 0$. However, the first excited state is degenerate and hence superpositions with non-zero velocity field are possible, yielding circles around the nucleus.

velocity field has the same dimension as it would in the spinless case. One could say that "spin" is a property of the wave function.

It is a simple calculation[7] to show that the velocity field of one silver atom (the wave function prepared as a Gaussian) after having passed through a Stern-Gerlach magnet is

$$v_Z^{\Psi}(z,t) = \frac{tz}{t^2+\mu^2} + \frac{\varphi}{m}\frac{\mu^2}{t^2+\mu^2}\tanh\left(\frac{2\mu\varphi tz}{t^2+\mu^2}\right)$$

where μ is the mass of the silver atom divided by the squared width (in Fourier space) of the initial Gaussian and φ is a constant that depends on all the magnetic properties of the system (it's of no interest to this discussion). One thing that is obvious is that this field is *not* constant. But this means that the trajectories are also not constant despite the lack of potential. One might be inclined to criticize the theory for such a result. But the important point is that Bohmian mechanics is not Newtonian and therefore does not have to obey the principles of Newtonian physics. The driving "force" one might seek due to imagining particles as newtonian particles comes from the wave function which behaves like a wave. This is what the guiding wave principle is: Waves guiding particles.

I previously mentioned that spin is not an intrinsic particle property. Since the only property a particle has is position, we need to be able to predict if a particle has spin up or down, using only that property and the initial conditions (initial position and wave function), if we want to show that it is sufficient to attribute the spin orientation solely to the wave function.

In order to see which initial positions lead to "particles with spin up/down", we need to take a closer look at the velocity field. The first property of the velocity field we need is that it is antisymmetric in z. The second property is that it is monotonically increasing/decreasing for positive/negative z. The velocity field along the trajectory is the first derivative of that curve. From this we can conclude that the integral curves of $v_z^{\Psi}, Z(t)$ above/below the x-axis are monotonically increasing/decreasing functions of time. In other words a particle with an initial position above/below the x-axis will be detected above/below the x-axis. Therefore a particle with an initial position above/below the x-axis will be considered a "spin up/down" particle.

[7]See Appendix.

So here we can explicitly see that an "observable" was reduced to the measurement of the position of the particle and that if given initial conditions, we can "predict" the outcome of an experiment using Bohmian mechanics. The subtlety that is highlighted by the quotation marks on the word "predict" is that we can't measure the initial position of the particle. It is this subtlety that explains why, despite having particles, we seem to always have a random element in our experiments, that cannot be dealt with by using clocks made in Switzerland.

10.5. References/Reading Material

- The book[4] is really a collection of papers on quantum mechanics by John Bell, mostly dealing with the Hidden Variables question. You find the famous Bell's Inequalities ("On the EPR paradoxon") paper in there as well. Although most of it is kind of old-fashioned (quantum mechanics of the 1960-1980's) and (mostly concerning Bohmian mechanics) not quite up-to-date, still worth reading.

- [20] is a basic introduction to Bohmian mechanics, but only available in German.

- [8] is THE book on Bohmian mechanics. If you want to know the whole story, and everything beyond. Contains as well a big chapter on probability theory and statistical physics that is worth reading even if you?re not that into Bohmian mechanics.

- The Homepage of Prof. Dürr's group [7]. See *Preprints* for a comprehensive list of papers on various issues in Bohmian mechanics and quantum mechanics.

We further want to mention the afterword of the standard quantum mechanics textbook by D. J. Griffihs[13]. It contains a very well presented overview on the EPR-Problem, the Hidden Variables question and even the discussion on non-locality vs. causality, especially the latter being worth reading.

10.6. Appendix

10.6.1. Derivation of the Stern-Gerlach velocity field

In order to simplify the calculation it is useful to perform a Galilei transformation on the system, such that we are in the rest-frame of the wave packet with respect to its x-velocity. We start with a wave function at rest (aside from spreading), turn on a Stern-Gerlach magnet for a short amount of time and calculate the wave function and trajectories after the magnet is turned off again. So we have to solve two Schrödinger equations:

- First we solve the Pauli equation to get the time evolution of a Gaussian in the magnetic field. This in particular yields a wave function with spin dependent phases.
- Then we take this solution again as an initial wave function and solve the free Schrödinger equation to get the spreading of the wave function after the magnet has been passed. From this solution we then calculate the velocity field of a particle after it has passed through a Stern-Gerlach magnet.

This is the basic idea, however the actual calculation is quite lengthy. For those who are interested, you will find the detailed calculation below.

Setup of the Stern-Gerlach thought experiment

For the initial wave function we will choose a spherically symmetric Gaussian, with a spinor in an eigenstate of σ_x. The wave function is chosen in this way because we are considering the case of a system that is symmetric with respect to the x-axis and because Gaussians are easy to Fourier transform. The spinor is chosen in such a way that neither eigenstate of σ_z is preferred in a Stern-Gerlach experiment. The wave function we will use is

$$\Psi(x,y,z) = \begin{pmatrix} \psi_+(z) \\ \psi_-(z) \end{pmatrix} \phi(x,y)$$

with

$$\phi(x,y) = \int dk_x \frac{1}{\sqrt{2\pi a}} e^{ik_x x - \frac{k_x^2}{2a}} \int dk_y \frac{1}{\sqrt{2\pi a}} e^{ik_y y - \frac{k_y^2}{2a}}$$

and

$$\psi_\pm(z) = \int dk_z \frac{1}{2\sqrt{\pi a}} e^{ik_z z - \frac{k_z^2}{2a}}$$

and $\sqrt{a}$ is the spread of the wave packet in k-space. $\psi_+(z)$ and $\psi_-(z)$ are the "up" and "down" components respectively.

In a Stern-Gerlach experiment the purpose of the magnet is to introduce a "spin" dependent phase, which causes the two wave packets to split. The magnetic field that would be required in order to change the wave function by nothing but a phase would only have a z-component. However such a magnetic field would have a divergence not equal to zero and would thus be forbidden by electrodynamics. In order to avoid this problem, the time the particle spends in the magnetic field is chosen to be short and the x and y components of the magnetic field are chosen such that they don't noticeably disturb $\phi(x, y)$. If these precautions are taken, the magnetic field effectively does nothing but introduce a phase.

The amount of time spent in the magnetic field will be called τ. τ is to be chosen such that it is much smaller than any time scale we encounter. The magnetic field will be chosen to be

$$\mathbf{B}(\mathbf{x}) = \begin{pmatrix} d(x,y) \\ h(x,y) \\ bz \end{pmatrix}$$

such that d and h have negligible influence on $\phi(x, y)$. b is a constant that contributes to the phase as we will see. With this in mind we can assume that $\phi(x, y)$ does nothing particularly interesting and we can focus on the part of the wave function that changes

$$\psi_\pm(z) = \int dk \frac{1}{2\sqrt{\pi a}} e^{ikz - \frac{k^2}{2a}}$$

For notational simplicity the indices of k have been dropped.

Wave function after the Stern-Gerlach magnet

Since τ is small compared to all timescales, spreading can be neglected, which means we can neglect the Laplace term in the Pauli equation. Since we are assuming that $\phi(x, y)$ stays constant we can replace $\Psi(\mathbf{x}, t)$ by $\begin{pmatrix} \psi_+(z) \\ \psi_-(z) \end{pmatrix}$ in the Pauli equation, leaving us with

$$i\partial_t \begin{pmatrix} \psi_+(z) \\ \psi_-(z) \end{pmatrix} = -\mu_M bz \sigma_z \begin{pmatrix} \psi_+(z) \\ \psi_-(z) \end{pmatrix} = -\mu_M bz \begin{pmatrix} \psi_+(z) \\ -\psi_-(z) \end{pmatrix}$$

In order to solve these equations we will rewrite them as

$$i\partial_t \psi_\pm(z) = \mp \mu_m b z \psi_\pm(z) \tag{10.3}$$

The next step is to use separation of variables

$$\frac{i}{\psi_\pm(z)} d\psi_\pm(z) = \mp \mu_m b z dt$$

Multiplying this equation by $-i$ and integrating it over the duration of the magnetic field gives us

$$\int_{\psi_\pm(z,0)}^{\psi_\pm(z,\tau)} \frac{i}{\psi_\pm(z)} d\psi_\pm(z) = \int_0^\tau \pm i \mu_M b z dt$$

If we integrate both sides and remember that $\psi_\pm(z,0) = \psi_\pm(z)$, we get

$$\ln\left(\frac{\psi_\pm(z,\tau)}{\psi_\pm(z)}\right) = \pm i \mu_M b z (\tau - 0)$$

Solving this equation for $\psi_\pm(z,\tau)$ gives us

$$\psi_\pm(z,\tau) = \psi_\pm(z) e^{\pm i \mu_M b z \tau} \tag{10.4}$$

If we write $\psi_\pm(z)$ in integral form, call the phase gained by the magnet $\varphi := \mu_M b \tau$ and pull the exponential into the integral, the final result is

$$\psi_\pm(z,\tau) = \int dk \frac{1}{2\sqrt{\pi a}} e^{i(kz \pm \varphi z) - \frac{k^2}{2a}} \tag{10.5}$$

Now that we have the wave function for the time the magnet is turned on, we can determine the wave function for the time afterward. In order to do this we will need to solve the Pauli equation for freely propagating waves, with the initial condition given by equation 10.5. The free Pauli equation in our case is

$$i\partial_t \psi_\pm(z,t) = \frac{\Delta}{2m} \psi_\pm(z,t) \tag{10.6}$$

The formal solution of this equation is

$$\psi_\pm(z,t) = \int dk \frac{1}{2\sqrt{\pi a}} e^{i(kz \pm \varphi z) - \frac{k^2}{2a}} e^{-i\omega_\pm(t-\tau)}$$

So what now remains is to determine $\omega_\pm$, which can be achieved by calculating

$$-\frac{\Delta}{2m}\psi_\pm(z) = -\frac{\Delta}{2m}\int dk \frac{1}{2\sqrt{\pi a}} e^{i(kz\pm\varphi z)-\frac{k^2}{2a}}$$

Because we're integrating over a Gaussian we can pull the derivatives into the integral leaving us with

$$-\int dk \frac{1}{2\sqrt{\pi a}} \frac{\Delta}{2m} e^{i(kz\pm\varphi z)-\frac{k^2}{2a}} = \int dk \frac{1}{2\sqrt{\pi a}} \frac{(k\pm\varphi)^2}{2m} e^{i(kz\pm\varphi z)-\frac{k^2}{2a}}$$

Inserting this result into equation 10.6 and computing the time derivative on the left hand side of the equation, leaves us with

$$\int dk \frac{\omega_\pm}{2\sqrt{\pi a}} e^{i(kz\pm\varphi z)-\frac{k^2}{2a}} = \int dk \frac{1}{2\sqrt{\pi a}} \frac{(k\pm\varphi)^2}{2m} e^{i(kz\pm\varphi z)-\frac{k^2}{2a}}$$

From this we can conclude that $\omega_\pm = \frac{(k\pm\varphi)^2}{2m}$. More generally we can conclude that for wave functions of the form

$$\psi(z) = \int dk\, e^{ig(k)z} l(k)$$

the time evolution in the free case is given by

$$\psi(z) = \int dk\, e^{ig(k)z} e^{-i\omega(k)(t-t_0)} l(k) \tag{10.7}$$

where $\omega = \frac{g(k)}{2m}$ and t_0 is the time for which the initial wave function holds. Since τ is much smaller than all timescales we will encounter, we will neglect it in the time evolution, leaving us with

$$\psi_\pm(z,t) = \int dk \frac{1}{2\sqrt{\pi a}} e^{i(kz\pm\varphi z-\omega_\pm t)-\frac{k^2}{2a}} \tag{10.8}$$

as the wave function for the time after the first magnet.

In order to see that equation 10.8 describes a wave function that is splitting, we need to calculate the integral. We will first sort the exponential in the

integrand in order to reduce the integration to a simple Gauss integral.

$$\begin{aligned} i(kz \pm \varphi z - \omega_{\pm} t) - \tfrac{k^2}{2a} &= i\left(kz \pm \varphi z - \frac{(k \pm \varphi)^2}{2m} t\right) - \tfrac{k^2}{2a} \\ &= i\left(kz \pm \varphi z - \frac{(k^2 \pm 2k\varphi + \varphi^2}{2m} t\right) - \tfrac{k^2}{2a} \\ &= -\frac{k^2}{2}\left(i\frac{t}{m} + \frac{1}{a}\right) + ik\left(z \mp \frac{\varphi t}{m}\right) + i\left(-\frac{\varphi^2 t}{2m} \pm \varphi z\right) \\ &= \frac{f}{2}\left(z \mp \frac{\varphi t}{m}\right)^2 \left(i\frac{t}{m} - \frac{1}{a}\right) - i\varphi\left(\frac{\varphi t}{m} \mp z\right) \\ &\quad - \left(i\frac{t}{2m} + \frac{1}{2a}\right)\left(k - i\frac{\left(z \mp \frac{\varphi t}{m}\right)}{\left(i\frac{t}{m} + \frac{1}{a}\right)}\right)^2 \end{aligned} \tag{10.9}$$

In the last step we introduced a time dependent function

$$f = \left(\frac{t^2}{m^2} + \frac{1}{a^2}\right)^{-1}$$

which will help shorten notation. Because f s not dependent on the wave packet (i.e., $\pm$), z or k, we can view it as a constant for the calculation (even when we calculate the velocity field). The first two terms of 10.9 are independent of k, which means all that is left to integrate is

$$Z(t) := \int dk \, \frac{1}{2\sqrt{\pi a}} e^{-\left(i\frac{t}{2m} + \frac{1}{2a}\right)\left(k - i\frac{\left(z \mp \frac{\varphi t}{m}\right)}{\left(i\frac{t}{m} + \frac{1}{a}\right)}\right)^2}$$

If we do the substitution

$$k \to k + i\frac{\left(z \mp \frac{\varphi t}{m}\right)}{\left(i\frac{t}{m} + \frac{1}{a}\right)}$$

we find that

$$Z(t) = \int dk \, \frac{1}{2\sqrt{\pi a}} e^{-\left(\frac{1}{2a} + \frac{it}{2m}\right)k^2}$$

From this consideration we can see that $Z(t)$ is a constant with respect to our calculation in the same sense as f. If we sort 10.9 by real and imaginary parts and then write out the full wave function we get

$$\psi_{\pm}(z,t) = \frac{Z(t)}{2\sqrt{\pi a}} e^{i\left(\frac{tf}{2m}\left(z \mp \frac{\varphi t}{m}\right)^2 - \frac{\varphi^2 t}{2m} \pm \varphi z\right)} e^{-\frac{f}{2a}\left(z \mp \frac{\varphi t}{m}\right)^2}$$

In this form we can see that the wave function is a Gaussian which is modulated by

$$\frac{Z(t)}{2\sqrt{\pi a}} e^{i\left(\frac{tf}{2m}\left(z\mp\frac{\varphi t}{m}\right)^2-\frac{\varphi^2 t}{2m}\pm\varphi z\right)}$$

and has it?s maximum at $z = \pm\frac{\varphi t}{m}$. From this we can conclude that the wave has indeed split and one packet is moving in the positive z-direction and the other in the negative z-direction linearly in time.

Velocity field after the Stern-Gerlach magnet

Now that we have the wave function, we can determine the velocity field, which is given by equation (10.2). The x, y-dependence of the wave function is given by $\Phi(x, y)$, which is roughly constant in time. This means that the x- and y-components of the velocity field are also roughly constant and won't give us any new insight. With that in mind we can focus on the z-component of the velocity field given by

$$v_z^{\Psi}(z,t) = m^{-1}\frac{\mathrm{Im}(\psi_+^*\partial_z\psi_+ + \psi_-^*\partial_z\psi_-)}{\psi_+^*\psi_+ + \psi_-^*\psi_-}(z,t) \tag{10.10}$$

where a star denotes complex conjugation. In order to calculate this we will need a systematic approach in order to avoid writing down terms that will cancel in the end. The first step is to calculate the derivative of the wave function.

$$\partial_z\psi_\pm = \psi_\pm\partial_z\left(i\left(\frac{tf}{2m}\left(z\mp\frac{\varphi t}{m}\right)^2-\frac{\varphi^2 t}{2m}\pm\varphi z\right)-\frac{f}{2a}\left(z\mp\frac{\varphi t}{m}\right)^2\right)$$

Before we go further it is important to note that the real part of the derivative of the exponential will not contribute to the velocity field, because $\psi_\pm^*\psi_p m$ is real. That leaves us with the contributing part of the derivative of the exponential of the wave function

$$\begin{aligned} i\partial_z\left(\frac{tf}{2m}\left(z\mp\frac{\varphi t}{m}\right)^2-\frac{\varphi^2 t}{2m}\pm\varphi z\right) &= i\left(\frac{tf}{m}\left(z\mp\frac{\varphi t}{m}\right)\pm\varphi\right) \\ &= i\left(\frac{tf}{m}z\pm\left(\varphi - f\frac{\varphi t^2}{m^2}\right)\right) \end{aligned}$$

In the last step we sorted the derivative by parts that are dependent on the wave packet and parts that are not. For the part that is not dependent on

the wave packet we can ??calculate

$$\frac{\mathrm{Im}(\psi_+^*\partial_z\psi_+ + \psi_-^*\partial_z\psi_-)}{\psi_+^*\psi_+ + \psi_-^*\psi_-} = \frac{tf}{m}z$$

because $\psi_+^*\psi_+ + \psi_-^*\psi_-$ is real. For the other parts we can identify the structure

$$\frac{\mathrm{Im}\left[i\left(\varphi - f\frac{\varphi t^2}{m^2}\right)(\psi_+^*\psi_+ - \psi_-^*\psi_-)\right]}{\psi_+^*\psi_+ + \psi_-^*\psi_-} = \left(\varphi - f\frac{\varphi t^2}{m^2}\right)\frac{\psi_+^*\psi_+ - \psi_-^*\psi_-}{\psi_+^*\psi_+ + \psi_-^*\psi_-}$$

The next step is to analyze

$$\frac{\psi_+^*\psi_+ - \psi_-^*\psi_-}{\psi_+^*\psi_+ + \psi_-^*\psi_-}$$

The first thing we can say is that all factors that ψ_+ and ψ_- have in common will cancel because they appear in the denominator and enumerator. Furthermore the imaginary part of the exponential of the wave function will vanish as well because we?re always multiplying components of the wave function by their complex conjugate (i.e., $\exp(iX)\times\exp(-iX) = 1$)). Taking these facts into consideration, (10.11) reduces to

$$\frac{e^{\frac{2f\varphi zt}{ma}} - e^{-\frac{2f\varphi zt}{ma}}}{e^{\frac{2f\varphi zt}{ma}} + e^{-\frac{2f\varphi zt}{ma}}} = \tanh\left(\frac{2f\varphi zt}{ma}\right)$$

If we collect our results the velocity field now reads

$$v_z^{\Psi}(z,t) = \frac{tf}{m^2}z + m^{-1}\left(\varphi - f\frac{\varphi t^2}{m^2}\right)\tanh\left(\frac{2f\varphi zt}{ma}\right)$$

Since we will want to analyze the behavior of the velocity field with respect to time we will write out f and denote $\frac{m}{a} := \mu$, leading us to the final result

$$v_z^{\Psi}(z,t) = \frac{tz}{t^2+\mu^2} + \frac{\varphi}{m}\frac{\mu^2}{t^2+\mu^2}\tanh\left(\frac{2\mu\varphi tz}{t^2+\mu^2}\right)$$

11. Bibliography

[1] SE Ahnert and MC Payne. General implementation of all possible positive-operator- value measurements of single-photon polarization states. Physical Review A, 71(1): 012330, 2005.

[2] Christian Bär and Christian Becker. C^*-algebras. In Christian Bär and Klaus Fredenhagen, editors, Quantum Field Theory on Curved Spacetimes, volume 786 of Lecture Notes in Physics, pages 1-37. Springer Berlin / Heidelberg, 2009. ISBN 978-3-642-02779-6. URL http://dx.doi.org/10.1007/978-3-642-02780-2_1. 10.1007/978-3-642-02780-2 1.

[3] J.S. Bell. On the einstein-podolsky-rosen paradox. Physics, 1(3):195-200, 1964. URL http://www.physics.princeton.edu/~mcdonald/examples/QM/bell_physics_\%1_195_64.pdf.

[4] J.S. Bell. Speakable and unspeakable in quantum mechanics: collected papers on quantum philosophy. Collected papers on quantum philosophy. Cambridge University Press, 2004. ISBN 9780521523387. LCCN 2004557644. URL http://books.google.de/books?id=FGnnHxh2YtQC.

[5] G. Benenti, G. Casati, and G. Strini. Principles of Quantum Computation and Information: Basic tools and special topics. Principles of Quantum Computation and Information. World Scientific Pub., 2007. ISBN 9789812565280. LCCN 2006278038. URL http://books.google.de/books?id=LpAkAQAAIAAJ.

[6] Sidney Coleman. Quantum mechanics in your face. video, April 1994. http://media.physics.harvard.edu/video/?id=SidneyColemanQMIYF.

[7] D Dürr. URL http://www.mathematik.uni-muenchen.de/\~bohmmech/BohmHome/bmstartE.htm .

[8] D. Dürr and S. Teufel. Bohmian mechanics: the physics and mathematics of quantum theory. Springer, 2009. ISBN 9783540893448. URL http://books.google.de/books?id=UWP2ZSs-UD0C.

[9] Einstein, A., Podolsky, B., and Rosen, N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? Physical Review, 47(10):777-780, May 1935. ISSN 0031-899X. URL http://link.aps.org/doi/10.1103/PhysRev.47.777.

[10] Klaus Fredenhagen. Algebraische Quantenfeldtheorie. lecture notes, 2003. URL http://www.itp.uni-hannover.de/~lechtenf/Events/Lectures/fredenhagen.pdf.

[11] M. Gell-Mann and J.B. Hartle. Quantum mechanics in the light of quantum cosmology. In S. Nakajima, Y. Murayama, and A. Tonomura, editors, Foundations of quantum mechanics in the light of new technology, volume 4 of Advanced Series in Applied Physics. World Scientific, 1990.

[12] Daniel M. Greenberger, Michael A. Horne, Abner Shimony, and Anton Zeilinger. Bell?s theorem without inequalities. American Journal of Physics, 58(12):1131-1143, 1990. URL http://link.aip.org/link/?AJP/58/1131/1.

[13] R.B. Griffihs. Consistent quantum theory. Cambridge University Press, 2003. ISBN 9780521539296. URL http://books.google.de/books?id=BzughgOxIiwC.

[14] S. Kochen and E.P. Specker. The problem of hidden variables in quantum mechanics. Indiana Univ. Math. J., 17(1):59-87, 1986. ISBN 9027705674

[15] N. David Mermin. Quantum mysteries revisited. American Journal of Physics, 58 (8):731-734, 1990. URL http://link.aip.org/link/?AJP/58/731/1.

[16] N. David Mermin. Simple unified form for the major no-hidden-variables theorems. Physical Review Letters, 65(27):3373-3376, December 1990. ISSN 0031-9007. URL http://link.aps.org/doi/10.1103/PhysRevLett.65.3373.

[17] N. David Mermin. Could Feynman Have Said This? Physics Today, 57(5):050000-12, May 2004.

[18] Michael A. Nielsen and Isaac L. Chuang. Quantum Computation and Quantum Information (Cambridge Series on Information and the Natural Sciences). Cambridge University Press, 1 edition, January 01, 2004. ISBN 0521635039. URL http://www.worldcat.org/isbn/521635039.

[19] H. Nikolic. Making nonlocal reality compatible with relativity. Arxiv preprint arXiv:1002.3226, 2010.

[20] O. Passon. Bohmsche Mechanik. Deutsch, 2004. ISBN 9783817117420. URL http://books.google.de/books?id=YrG2vn1n6zUC.

[21] John Preskill. Quantum computation. Lecture Notes. URL http://www.theory.caltech.edu/people/preskill/ph229/notes/chap3.pdf.

[22] S.J. Summers. John von Neumann and the foundations of quantum physics, volume 8, chapter On the Stone-von Neumann Uniqueness Theorem and Its Ramifications, page 135. Kluwer Academic Pub, 2001.

[23] John von Neumann. Die Eindeutigkeit der Schrödingerschen Operatoren. Math. Ann., 104:570=578, 1931.