Contextuality in a Deterministic Quantum Theory

Atul Singh Arora

Department of Physical Sciences

Indian Institute of Science Education and Research, Mohali

This dissertation is submitted for the degree of

Master of Science

2016

I would like to dedicate this thesis to my mother, who introduced me to the most valuable notion, that of freedom.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text.

Atul Singh Arora
2016

Acknowledgements

I acknowledge the contribution of my project guide, Prof. Arvind, who provided me with an environment conducive to research, and motivation to facilitate completion of this project. He encouraged independent thinking and guided me through difficulties I encountered.

Discussions with the Quantum Computation & Quantum Information (QCQI) group members, especially those with Rajendra Bhati and Kishor Bharti have been particularly efficacious. Inputs from Jaskaran Singh have also been helpful. The QCQI group seminars, specifically, those by Samridhi Gambhir and Dr. Arun Shehrawat, have helped narrowing the thesis problem.

Dr. Abhishek Choudhuri and Prof. Sudeshna Sinha, although formally not related to the project, have been kind enough to assist me at various instances with challenges I faced while numerically simulating the equations.

I am grateful to my parents and my colleagues/friends, who have been pivotal to helping me maintain mental balance, inspired me and provided emotional support, in ways they perhaps don't realize, specifically Manu Jayadharan, Prashansa Gupta, Kishor Bharti, Ritu Roy Choudhuri, Vivek Sagar, Yosman Bapat-dhar, Saumya Gupta, Shwetha Srinivasan, Evelyn Abraham and Srijit Mukherjee.

I end with acknowledging Immanuel Kant, for his philosophical work on ethics, which has heavily affected my thoughts, actions and outlook to life.

Abstract

The Copenhagen Interpretation of Quantum Mechanics (QM) asserts that the wavefunction is the most complete description, which entails that there is an inherent fuzziness in our description of nature. There exists a completion of QM, known as Bohmian Mechanics (BM), which replaces this fuzziness with precision, and re-introduces notions of physical trajectories. Various interesting questions arise, solely by existence of such a description; doesn't it contradict the uncertainty principle, for instance. Most of these questions were found to have been addressed satisfactorily in the literature. There was however, one question, whose answer has become the subject of the thesis; that of the paradoxical co-existence of contextuality and BM. Contextuality arguments show that the value an operator takes, must depend, in addition to the state of the system (including any hidden variables), on the complete set of compatible operators, to be consistent with QM. BM being deterministic, is at complete odds with this notion.

We were able to show, after various attempts, that the notion of contextuality is infact not necessary, by identifying another 'classical property' and constructing a non-contextual toy-model, serving as a counter-example to the impossibility proof.

Contents

List of Figures

List of Tables

Chapter 1

Prologue

1.1 Overview

A course in Quantum Mechanics (QM) typically leaves the reader with a fuzzy view of the world. The founders of the subject, were able to abstract out the mathematics from it's implications and interpretation. The view that the wavefunction is the most complete description of a system, known as the 'Orthodox Copenhagen Interpretation', is among the most popular. Einstein and Bohm, both played an important role in challenging this belief.
Einstein showed that if one makes reasonable assumptions about nature (apparently these were in spirit with his relativity theories), then QM must be incomplete. By incompleteness, he meant that the results of certain experiments should be predictable precisely, but QM fails at doing this. Thus he concluded that, QM must be treated as an intermediate theory and that a more complete description of nature should be possible. We will come back to this discussion. Bohm challenged this view by realizing that one can't experimentally refute the interpretation. His argument was that if for instance, the predictions of QM fail to match with experiments, then one can always add appropriate terms in the Hamiltonian until the difficulty is resolved. He was already anticipating the modern form of particle physics, where to accurately account for interactions between particles, one can postulate new force mediating particles, which was also done in the history of the subject, but can't refute the interpretation due to inaccuracy in predictions. He therefore, aimed at, and succeeded at constructing a complete quantum theory, now known as Bohmian Mechanics (BM), existence of which, would show that we are not justified at believing the interpretation, simply because there exist other alternatives.

Returning to Einstein, as will be described in more detail, Bell was able to construct a physical situation, that would test the very requirements Einstein imposed on a complete quantum theory; he was able to show that according to predictions of QM, such a complete description is impossible. This was verified experimentally (and was confirmed to be true, without any

possible loop-holes, only recently in 2015).
It might appear that Bell's test shows that QM is inconsistent with BM, but upon looking at the details, one can be convinced, that this is not the case. Infact, Bell was among the first few to popularize BM.

The discussion so far, is well known in the literature. However, certain developments which followed the said work of Bohm, again lead to apparent inconsistencies between BM and QM, which hadn't been satisfactorily addressed in the literature. These are the subject of the thesis.
For the sake of completeness, we merely name these, for one must look at the details, to be able to appreciate them. Greenberger, Horne and Zeilinger (GHZ), constructed a test that apparently showed determinism, the notion that observables have pre-defined values, can't exist. Developments due to Gleason, Bell himself, Kochen, Specker, Peres and Mermin, showed that the value an operator takes, must depend on the context in which it is measured; context here refers to the complete set of compatible operators.
The apparent contradiction must now be clear; BM is deterministic and it is at odds with the notion of contextuality as well.

1.2 The EPR argument

The EPR argument requires an entangled state, over two particles, which was originally written as

$$\psi(x_1,x_2) = \int e^{-i(x_1-x_2+x_0)p/\hbar}dp = \delta(x_1-x_2+x_0).$$

In the modern notation, one can write this as

$$\begin{aligned} |\psi\rangle &= \int \delta(x_1-x_2+x_0)dx_1dx_2\,|x_1\rangle_1\,|x_2\rangle_2 \\ &= \int |x\rangle_1\,|x+x_0\rangle_2\,dx. \end{aligned}$$

Similarly, one can write

$$\begin{aligned} |\psi\rangle &= \int e^{-i(x_1-x_2)p/\hbar}dpdx_1dx_2\,|x_1\rangle_1\,|x_2\rangle_2 \\ &= \int e^{-ix_1p/\hbar}\,|x_1\rangle_1\,dx_1e^{ix_2p/\hbar}\,|x_2\rangle_2\,dx_2dp \\ &= \int |p\rangle_1\,|-p\rangle_2\,dp. \end{aligned}$$

where the first ket is that of the first particle, and second that of the second. An attempt to reproduce the precise argument, will not be made. Instead, we satisfy ourselves with the

following simplified argument. We assume the principle of locality holds, which is to say that if two particles are sufficiently far away, then any change made to one particle, should not influence the other instantaneously. Given this assumption, consider two particles, which are sufficiently far away and that their state is given by $|\psi\rangle$. Now if the momentum of particle 2 is measured by observer 2, then observer 2, according to QM, will be able to predict the momentum of particle 1. Similarly, if the position of particle 2 is measured, then observer 2 can predict the position of particle 1. However, since the choice made by observer 2, can't influence particle 1 (by the assumption of locality), it follows that particle 1 had both it's position and momentum well defined, without being measured. Since QM fails to yield the answer with precision, whereas as we have shown, the answer was precise, we are forced to conclude that QM is not a complete description of nature.

Einstein ended his paper with this remark: "While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, that such a theory is possible."

1.3 Bell's Theorem

Bell's work addresses and satisfactorily answers Einstein's open question, of whether a *such* a description exists. Consider the following scenario. There are two observers, A and B, each has a particle. Each observer, can measure two properties, call them $\hat{a}_1, \hat{a}_2$ for observer A and $\hat{b}_1, \hat{b}_2$ for observer B, which yield a ± 1 outcome. Let $\langle \hat{a}_i \hat{b}_j \rangle$ represent the average value obtained by measuring the property $\hat{a}_i$ and $\hat{b}_j$ on the first and second particle respectively. Consider the average given by $\langle \hat{B} \rangle = \langle \hat{a}_1 \hat{b}_1 \rangle + \langle \hat{a}_1 \hat{b}_2 \rangle + \langle \hat{a}_2 \hat{b}_1 \rangle - \langle \hat{a}_2 \hat{b}_2 \rangle$. If one assumes local realism[1], that is that the values a_i takes are unaffected by those taken by b_i and that $\hat{a}_i, \hat{b}_j$ have pre-defined values (respectively), then $\langle \hat{B} \rangle \leq 2$. One can check this quickly by a brute force substitution of ± 1 values, in place of a_i, b_j.

Consider now, the state $|\psi\rangle = (|+-\rangle - |-+\rangle)/\sqrt{2}$, where $\hat{\sigma}_x|\pm\rangle = \pm|\pm\rangle$, and $\hat{\sigma}_{x,y,z}$ are the Pauli matrices. Let $\hat{a}_1 = \hat{\sigma}_z$ and $\hat{a}_2 = \hat{\sigma}_x$, while $\hat{b}_1 = -\frac{\hat{\sigma}_z + \hat{\sigma}_x}{\sqrt{2}}$ and $\hat{b}_2 = \frac{\hat{\sigma}_z - \hat{\sigma}_x}{\sqrt{2}}$. Upon evaluating $\langle \hat{B} \rangle$ using the rules of QM, one obtains $\langle \hat{B} \rangle = 2\sqrt{2} \nleq 2$. This entails that our assumption, that of local realism, must be incorrect, which settles the question: A *local* hidden variable description can not exist. The word hidden is used to represent information about the state of the system, that is not contained in the wavefunction. It is instructive to emphasize that a non-local description, may still be possible. It is also important to note that, despite this non-locality, one can't use QM to send signals instantaneously (superluminal communication

[1] Infact, an argument similar to EPR can be used to show that locality entails realism.

is not possible). This is known as the no-signalling theorem and can be proven by showing that marginals in QM, don't depend on the system which is traced out (TODO).

1.4 Bohm's Theory, Bohmian Mechanics (condensed introduction)

Bohm gave a precise, but non-local description of quantum phenomena. Let us start with non-interacting particles. A particle is associated with (1) a position, q & momentum, p, precisely defined and (2) a wavefunction $\psi = Re^{iS/\hbar}$. The postulates of the theory are: (a) Evolution of the wavefunction, is governed by Schrodinger's equation: $i\hbar\partial\psi/\partial t = -(\hbar^2/2m)\nabla^2\psi + V\psi$. (b) The particle is guided by the wavefunction: $\dot{q} = p/m$ where $p = \nabla S = \hbar\mathrm{Im}(\nabla\psi/\psi)$. (c) The initial distribution of the particles is given by $\rho(x) = |\psi|^2$.

The astute reader would've noticed that ∇S is just the probability current, which entails that if the initial distribution satisfies $|\psi(t_0)|^2$, then it will do so at all times t, thereafter. Before generalizing this to multiple particles, let us first see a quick consequence of this formulation. In the Orthodox Copenhagen interpretation, the double slit experiment is a source of mystery and the which slit question, that of confusion. In BM, since trajectories are well defined, one observes (TODO: add image) that the particle goes through precisely one slit and then later, forms the interference pattern. The wavefunction goes through both slits, and interferes to create the pattern. If one of the slits is blocked, then as expected, since the wavefunction can't interfere, the pattern is lost. We haven't spoken about measuring the particle yet, but if one notes that to measure the particle, the potential at one of the slits must be changed, then it is immediate that the interference pattern will be effected, since the wavefunction is affected by the potential at both slits, even if the particle passes through only one. One finds that by using the following multi-particle generalization, and an appropriate measuring scheme, one can show more precisely how the process of measuring effectively destroys the pattern, replacing the mystery with clarity.

For N interacting particles, we have $p_i = \nabla_i S(q_1, q_2, \ldots, q_N)$ where note that the momentum of the i^{th} particle, depends on the instantaneous positions of all the particles. Consequently, BM is an explicitly non-local, but complete description.

(TODO: change the x axis to z, and change explicit reference to x,y,z in favour of q_x, q_y, q_z etc.) As a final remark, it must be added that spins can also be included in BM, however, a particle is not associated with a specific spin, only it's wavefunction is. For a spinor, say $\psi \equiv (\psi_+, \psi_-)^T$, the generalization is that $p = \hbar\mathrm{Im}((\psi, \nabla\psi)/(\psi, \psi))$ where $(.,.)$ represents inner product in the spin space $\mathbb{C}^2$. A quick illustration is in order. Consider the following

Stern-Gerlach setup; Quantum effects are considered only along the x-axis. A particle moves along the z-axis, with speed v_z, and it's initial wavefunction, given by a Gaussian centred at the origin, say $\psi(x) = (1/\sqrt{2\pi}\sigma)e^{-x^2/2\sigma^2}$, viz. $|\psi\rangle = \int dx\psi(x)\,|x\rangle$. It's spin state is given by $|\chi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, where $|0\rangle$ and $|1\rangle$ represent computational basis. Along the z-axis, a strong heterogeneous magnetic field is present, whose action maybe captured by $H_{\text{int}} = a\hat{p}_x \otimes \hat{\sigma}_z$ where a is a constant that quantifies the strength of the field. Why this particular form works, will become clear momentarily. Assuming that a is large enough to neglect effects of free-evolution, we have $\hat{U}(t) = e^{-ia\hat{p}_x \otimes \hat{\sigma}_z t/\hbar}$. Thus, if the initial state is $|\Psi\rangle = |\psi\rangle \otimes |\chi\rangle$, then

$$\begin{aligned} |\Psi(t)\rangle &= \hat{U}(t)\,|\Psi\rangle \\ &= \frac{e^{-ia\hat{p}_x t/\hbar}\,|\psi\rangle \otimes |0\rangle + e^{ia\hat{p}_x t/\hbar}\,|\psi\rangle \otimes |1\rangle}{\sqrt{2}} \\ &= \frac{|\psi_{at}\rangle \otimes |0\rangle + |\psi_{-at}\rangle \otimes |1\rangle}{\sqrt{2}} \end{aligned}$$

where $|\psi_{x_0}\rangle \equiv \int dx\psi(x-x_0)\,|x\rangle$, viz. a Gaussian wavepacket centered at x_0. One can plot $|\Psi|^2$, as a function of z, using $z = v_z t$, schematically as shown in the figure (TODO: put the figure), where regions enclosing, say 70% of $|\Psi|^2$ have been outlined. So far, according to QM, if we measure the position of the particle, then (given $\sigma \ll at$), obtaining the particle near $x = at$, is as probably as finding it near $x = -at$. QM doesn't make a deterministic prediction at this stage. According to BM, if in addition to the wavefunction, we assume that the particle was initially at some $x > 0$, then we can predict precisely where the particle will turn up. Infact, one can accomplish this with essentially no calculation. Since $\dot{x} = \nabla S/m$, a single valued function, it follows that trajectories of particles won't intersect, where recall that S is given by $\psi = Re^{iS/\hbar}$. Consequently, a particle that starts with $x > 0$, must follow the 'up' trajectory, else by symmetry, the trajectories will have to intersect. We conclude therefore, that if $x > 0$ initially, the particle has spin $|0\rangle$ and if $x < 0$, the spin must be $|1\rangle$. This appears very clear and intuitive. A non-intuitive aspect of this, is that we can't associate spins with particles, even though we can predict precisely, the outcome of the experiment. This is manifested by the observation that if $H_{\text{int}} \to -H_{\text{int}}$, which practically amounts to reversing the heterogeneity of the magnet, then $\sqrt{2}\,|\Psi(t)\rangle = |\psi_{-at}\rangle \otimes |0\rangle + |\psi_{at}\rangle \otimes |1\rangle$. Again, if $x > 0$ initially for a particle, it will follow the 'up' trajectory. However, now this corresponds to spin $|1\rangle$, as opposed spin $|0\rangle$. We have thus demonstrated that we can't associate spin uniquely to a particle, and that it must only be associated with the wavefunction.

As an aside, it may be added that BM is known for removing the fundamental role of an observer from the description of QM. This is accomplished, roughly speaking, by introducing the positions of particles as well defined, and then explaining the 'collapse' of wavefunction,

by means of the particle's interaction with those in the environment, which entails that the 'collapsed wavefunction' serves as an effective wavefunction, which can be used to describe the motion of the particle henceforth.

1.5 Determinism: The GHZ test

One drawback of Bell's test, was that it's statistical nature. The GHZ test, takes this a step further and shows that one can't even conceive of having pre-defined values. The construction requires three observers with one particles each. Each observer can measure two properties, X and Y, with outcomes ± 1. We start with the state $\sqrt{2}\,|\chi_G\rangle = |000\rangle - |111\rangle$ and note that for $\hat{A} := \hat{\sigma}_x \otimes \hat{\sigma}_y \otimes \hat{\sigma}_y$, $\hat{A}\,|\chi_G\rangle = |\chi_G\rangle$, where the property X is the projection of spin of the particle along the x-axis and Y is that along the y-axis. Thus a measurement of $\hat{A}$ (the first observer measures X and the other measure Y), will yield $+1$ with certainty. Next, we define $\hat{B} := \hat{\sigma}_y \otimes \hat{\sigma}_x \otimes \hat{\sigma}_y$ and $\hat{C} := \hat{\sigma}_y \otimes \hat{\sigma}_y \otimes \hat{\sigma}_z$, which by symmetry, must also yield $+1$ for the said state. If these observables, X and Y had predefined values, then a measurement of $\hat{A}\hat{B}\hat{C}$ would be equivalent to measuring $\hat{D} := \hat{\sigma}_x \otimes \hat{\sigma}_x \otimes \hat{\sigma}_x$, since Ys appear twice for each particle, and $Y^2 = 1$. Since a measurement of each $\hat{A}$, $\hat{B}$ and $\hat{C}$ yields a $+1$, it follows that a measurement of $\hat{D}$ therefore, must also yield $+1$. However, $\hat{D}\,|\chi_G\rangle = -\,|\chi_G\rangle$, which entails that a measurement of $\hat{D}$ must yield a -1. Thus, we arrive at a contradiction and conclude that the assumption that the properties X and Y had predefined values, must be incorrect.

1.6 Contextuality: The Peres Mermin test

There've been various developments, of which the simplest, with effectively the same consequence, is discussed: the Peres Mermin test. In this test also, we will show that pre-defined values can't exist, but in addition, also define the notion of contextuality. Consider the following set of operators

$$\hat{A}_{ij} \doteq \begin{bmatrix} \hat{\mathbb{I}} \otimes \hat{\sigma}_x & \hat{\sigma}_x \otimes \hat{\mathbb{I}} & \hat{\sigma}_x \otimes \hat{\sigma}_x \\ \hat{\sigma}_y \otimes \hat{\mathbb{I}} & \hat{\mathbb{I}} \otimes \hat{\sigma}_y & \hat{\sigma}_y \otimes \hat{\sigma}_y \\ \hat{\sigma}_y \otimes \hat{\sigma}_x & \hat{\sigma}_x \otimes \hat{\sigma}_y & \hat{\sigma}_z \otimes \hat{\sigma}_z \end{bmatrix}$$

which have the peculiar property that all operators along a row (and along a column) commute. It is trivial to see that this holds for the first two rows and the first two columns. To see that this holds also for the last row and column, note the anti-commutation relation, $\{\hat{\sigma}_x, \hat{\sigma}_y\} = 0$ and that $\sigma_z = i\sigma_y\sigma_x$, which one may check explicitly. Another interesting property is that the product of rows (columns) yield $\hat{R}_i = \mathbb{I}$ and $\hat{C}_j = \mathbb{I}\,(j \neq 3)$, $\hat{C}_3 = -\mathbb{I}$, $(\forall\, i, j)$ where $\hat{R}_i \equiv \prod_j \hat{A}_{ij}$, $\hat{C}_j \equiv \prod_i \hat{A}_{ij}$. This can be verified easily by using the aforesaid relations and the fact that $\sigma^2 = 1$,

for every Pauli matrix. Using this property of R_i and C_j, it is easy to show that no pre-defined values for operators can exist. Let us assume that pre-defined values do exist. Note that to get $C_3 = -1$, we must have an odd number of -1 assignments in the third column. In the remaining columns, the number of -1 assignments must be even for each column. Thus, in the entire square, the number of -1 assignments must be odd. Let us use the same reasoning, but along the rows. Since each $R_i = 1$, we must have even number of -1 assignments along each row. Thus, in the entire square, the number of -1 assignments must be even. We have arrived at a contradiction and therefore we conclude that our assumption that operators have predefined values, must be wrong. Note that unlike the GHZ test, the Peres Mermin test is state independent.

We are now in a position to discuss the notion of contextuality. We begin with defining a non-contextual assignment to be one, where the assignment depends only on the state ($|\psi\rangle$ + hidden variables, if any) and the operator to which the assignment is being made. It follows that we had tacitly assumed a non-contextual assignment, that resulted in a contradiction. If we allow for the value of an operator, to depend on the context in which it is measured, where the context is meant to refer to a set of compatible observables, then the contradiction won't arise. For instance, if we imagine that all observables yield a $+1$, except $\hat{A}_{33}$ which yields a -1 if measured with $\hat{A}_{32}, \hat{A}_{31}$ and yields a $+1$ if measured with $\hat{A}_{23}, \hat{A}_{13}$. If this doesn't appear entirely unsatisfactory, then you're on the right track and are encouraged to read the more detailed description given in Section [??].

Chapter 2

Determinism Tests and Bohmian Mechanics

2.1 Rationale

As was pointed out in the previous chapter, it is known that BM doesn't assign a unique value to spin of a particle. For BM, spin is only a property of the wavefunction. Recall that the GHZ test was formulated for spins. Consequently, the conclusion that spins can't have pre-defined values (deterministic) does not contradict BM outright. Precisely how the GHZ test is compatible with BM, is worth exploring, for it will edify our understanding of how to analyse a physical situation using BM and the relation between determinism and non-locality.

From the point of view of QM however, the treatment of spins is not fundamentally different from that of phase-space variables (position q and momentum p). It is not surprising therefore, that the GHZ test can be generalized to phase space and at least one such extension is known. The conclusion one draws from the phase space GHZ test, would then be that q, p can't have pre-defined values. This then, is in direct contradiction with BM, which claims that q, p are precisely defined and their evolution completely determined, given the initial position, q_0, and the wavefunction. Since it is believed that BM is completely consistent with QM, it is of considerable interest to explore how BM can resolve the apparent contradiction, paradox if you will. If it fails, then we would have identified a way to falsify BM.

2.2 GHZ

For convenience, we recall that in our previous discussion, $\hat{A} \equiv \hat{\sigma}_x \otimes \hat{\sigma}_y \otimes \hat{\sigma}_y$, $\hat{B} \equiv \hat{\sigma}_y \otimes \hat{\sigma}_x \otimes \hat{\sigma}_y$, $\hat{C} \equiv \hat{\sigma}_y \otimes \hat{\sigma}_y \otimes \hat{\sigma}_x$, and $\hat{D} \equiv \hat{\sigma}_x \otimes \hat{\sigma}_x \otimes \hat{\sigma}_x$, which are such that $\hat{A}\,|\chi\rangle = \hat{B}\,|\chi\rangle = \hat{C}\,|\chi\rangle = |\chi\rangle$, while

$\hat{D}|\chi\rangle = -|\chi\rangle$, for $\sqrt{2}|\chi\rangle = |000\rangle - |111\rangle$.

2.2.1 Compatibility with BM

To analyse any situation using BM, one is required to know the experimental arrangement. In this case, we assume that spins are measured using the Stern-Gerlach (SG) apparatus. Without loss of generality, let us assume that the initial state of the particle is given by, $\sqrt{2}|\Psi(\vec{r}_1,\vec{r}_2,\vec{r}_3,t=0)\rangle = |\psi_{+++}(\vec{r}_1,\vec{r}_2,\vec{r}_3,t=0)\rangle \otimes |000\rangle - |\psi_{---}(\vec{r}_1,\vec{r}_2,\vec{r}_3,t=0)\rangle \otimes |111\rangle$, where $\vec{r}_i$ represents the position vector in the frame of the i^{th} observer. Note that the explicit tensor product is used to separate the spin and position parts of all three particles. If the wavefunction for each particle is assumed to be Gaussian initially and propogating along their respective axes of the SG apparatus, then one can further simplify the form of $|\psi_{\pm\pm\pm}\rangle$. It has been shown [??] that the time evolution of $|\psi_{\pm\pm\pm}\rangle$ can be written as products of 3 single particle solutions of the SG setup, which was analyzed by Bohm himself. Once $|\Psi(\vec{r}_i,t)\rangle$ is known, one can evaluate the equation of motion for the three particles, using BM. If the SG apparatus are setup to measure say XYY, then from both numerical simulations & analysis of the trajectory equations, it is observed that (in the direction relevant to measurement), four attractor basins are formed: $(+++)$, $(+--)$, $(-+-)$, and $(--+)$, where $\pm$ represent the physical location in the SG, corresponding to a spin 'up' ('down') measurement. The product is always $+1$, consistent with predictions of QM. However, when the SG apparatus are setup to measure XXX, the trajectories are found to obey equations which possess four attractive basins: $(---)$, $(-++)$, $(+-+)$ and $(++-)$. Again, the product is -1, in agreement with QM.

As a remark, it maybe stated, that the details of evaluating the trajectory, become complicated rather quickly and that it is unlikely to obtain analytic solutions for the phase-space scenario trivially. Consequently, numerical simulations are a necessity beyond this stage. [TODO: add details of the calculations?]

In conclusion, one finds that non-locality enters the description from the fact that the attractor basins which form, depend on the settings of *all* SG apparatus. Thus, while all the results are deterministic, we learn that the results depend on the precise experimental setup, and not merely operators.

2.2.2 Phase Space

To extend the GHZ test to phase space, note first that we need only the following situation. Consider instead of observables (which are Hermitian), unitary operators, $\hat{X}$ and $\hat{Y}$ with the following redefinitions: $\hat{A} \equiv \hat{X}^\dagger \otimes \hat{Y} \otimes \hat{Y}^\dagger$, $\hat{B} \equiv \hat{Y}^\dagger \otimes \hat{X}^\dagger \otimes \hat{Y}$, $\hat{C} \equiv \hat{Y} \otimes \hat{Y}^\dagger \otimes \hat{X}^\dagger$ and $D = \hat{X} \otimes$

$\hat{X}\otimes\hat{X}$. If these unitary operators also satisfy $\{X,Y\}=0=\{X,Y^\dagger\}$, then it follows that (a) $\hat{A}$, $\hat{B}$, $\hat{C}$ and $\hat{D}$ all commute and (b) $\hat{A}\hat{B}\hat{C}\hat{D}=-\mathbb{I}$. Now any simultaneous eigenket of $\hat{A}$, $\hat{B}$, $\hat{C}$ and $\hat{D}$ will yield a GHZ situation.

To see why that will work, replace the unitary operators with complex numbers and note that the unitary property translates to each of these numbers being uni-modular. Thus, $ABC = X^*\otimes X^*\otimes X^*$, using $Y^*Y=1$ for each particle. This would entail that $ABC=D^*$, viz. $ABCD=1$. However we also know that $ABCD=-1$, which yields the contradiction.

Another patent issue with this scheme, is the use of unitary operators, as opposed to observables. This issue is resolved by explicit construction, however the idea can be stated in general. If an arbitrary unitary operator, is a function of some fixed observable, and that alone, then one can measure the said observable to evaluate the value of the unitary operator, thereby dissolving the objection.

2.2.2.1 Known Extension

One possible construction[??], involves the use appropriate displacement operators. $\hat{X}\equiv e^{i\sqrt{\pi}\hat{q}/L}$ and $\hat{Y}\equiv e^{i\sqrt{\pi}\hat{p}L}$, where L is some length scale and units are s.t. $\hbar=1$ (for this section). These satisfy $\{\hat{X},\hat{Y}\}=0$, which follows trivially by recalling that $e^{i\hat{p}u}e^{i\hat{q}v}=e^{i\overset{1}{\hbar}uv}e^{i\hat{q}v}e^{i\hat{p}u}$. To construct a simultaneous eigenstate, observe that for

$$\begin{aligned}
\sqrt{2}\,|\uparrow\rangle_{q_0,p_0} &\equiv \sum_{k=-\infty}^{\infty} e^{i\sqrt{\pi}2kp_0L}\left|q=\sqrt{\pi}L(q_0+2k)\right\rangle \\
&\quad +i\sum_{k=-\infty}^{\infty} e^{i\sqrt{\pi}(2k+1)p_0L}\left|q=\sqrt{\pi}L(q_0+2k+1)\right\rangle, \\
\sqrt{2}\,|\downarrow\rangle_{q_0,p_0} &\equiv \sum_{k=-\infty}^{\infty} e^{i\sqrt{\pi}2kp_0L}\left|q=\sqrt{\pi}L(q_0+2k)\right\rangle \\
&\quad -i\sum_{k=-\infty}^{\infty} e^{i\sqrt{\pi}(2k+1)p_0L}\left|q=\sqrt{\pi}L(q_0+2k+1)\right\rangle,
\end{aligned}$$

where $p_0L/\sqrt{\pi}$ & $\sqrt{\pi}Lq_0\in[0,1)$, $\hat{X}\,|\uparrow\rangle=|\downarrow\rangle$, $\hat{Y}\,|\uparrow\rangle=i\,|\downarrow\rangle$ and similarly $\hat{X}\,|\downarrow\rangle=|\uparrow\rangle$, $\hat{Y}\,|\downarrow\rangle=-i\,|\uparrow\rangle$ (we have dropped q_0 and p_0 for simplicity). This also holds for $\hat{X}^\dagger$ and $\hat{Y}^\dagger$. To see this, note that $\hat{Y}^\dagger\,|q\rangle=\left|q+\sqrt{\pi}L\right\rangle$ while $\hat{X}\,|x\rangle=e^{i\sqrt{\pi}q/L}\,|q\rangle$. If one defines $\hat{Z}=i\hat{Y}\hat{X}$, from the aforesaid, it follows $\hat{Z}\,|\uparrow\rangle=|\uparrow\rangle$ and $\hat{Z}\,|\downarrow\rangle=-\,|\downarrow\rangle$. The problem has been made sufficiently analogous to the original GHZ test. It is now immediate that the required simultaneous entangled eigenket must be

$$|\psi\rangle=\frac{|\uparrow\uparrow\uparrow\rangle-|\downarrow\downarrow\downarrow\rangle}{\sqrt{2}}.$$

It is obvious that for obtaining the value of $\hat{X}$, one need only measure $\hat{q}$ and $\hat{p}$ to obtain the value of $\hat{Y}$ and their conjugates. We have therefore an extension of the GHZ test to continuous variables. To understand how BM explains this however, one must be able to simulate this. The test in the given form, is not simple to simulate, since the states involved have infinite spread in position space (infact one can show that even in momentum space the spread is infinite). Further, the wavefunction in position space, is a countable union of disjoint delta functions. Neither of these are desired from the numerical point of view.

2.2.2.2 Optimized Extension

This result was obtained after considerable effort. We construct an optimization of the phase space GHZ test, such that (a) the wavefunctions aren't sharp (no delta functions) and (b) that they disappear as $q \to \pm\infty$. First of all, consider the same states $|\psi_0\rangle, |\psi_1\rangle$ for $N = 8$, as those considered in my first paper (TODO: define these properly). Recall that for $\hat{Z}$ as defined there, viz. $\hat{Z} = Z(\hat{q} \mod 2L)$, or more precisely as $\hat{Z} = Z(\hat{q}) = \mathrm{sgn}(\sin(\hat{q}\pi/L))$, we had $\hat{Z}|\psi_0\rangle = |\psi_0\rangle$ and $\hat{Z}|\psi_1\rangle = -|\psi_1\rangle$. In addition to this, I define $\hat{X} = e^{-i\hat{p}L/\hbar}$ (as opposed to defining it to be hermitian). Now, we know that$|\psi_\pm\rangle \equiv \frac{|\psi_0\rangle + |\psi_1\rangle}{\sqrt{2}}$ is not an eigenstate of $\hat{X}$. So we optimize the observable $\hat{X}$ to $\hat{X}' \equiv \hat{X}\hat{T}$, where $\hat{T} \equiv e^{i\hat{p}NLa(\hat{q})/2}$ where

$$a(q) = \begin{cases} 1 & 2L < q < 4L \\ 0 & \text{else} \end{cases}.$$

The idea is that you shift certain peaks to the right place, before applying the displacement operator $\hat{X}$. To illustrate this, consider explicitly $|\psi_0\rangle = (|\varphi_{-4}\rangle + |\varphi_{-2}\rangle + |\varphi_{-1}\rangle + |\varphi_{-3}\rangle)/\sqrt{4}$. The operation of $\hat{T}$ is $\hat{T}|\varphi_4\rangle = |\varphi_{-5}\rangle$, $\hat{T}|\varphi_3\rangle = |\varphi_{-6}\rangle$ and $\hat{T}|\varphi_n\rangle = |\varphi_n\rangle$ for $n \in \{-4,-3,-2,-1,1,2\}$. It is now evident that $\hat{X}'|\psi_0\rangle = \hat{X}\hat{T}|\psi_0\rangle = |\psi_1\rangle$. Note also that $\hat{X}'^\dagger|\psi_0\rangle = |\psi_1\rangle$. Similarly $\hat{X}'|\psi_1\rangle = |\psi_0\rangle$ and $\hat{X}'^\dagger$ does the same. So finally, now consider $|\psi_G\rangle \equiv (|\psi_0\psi_0\psi_0\rangle - |\psi_1\psi_1\psi_1\rangle)/\sqrt{2}$. With $\hat{A} \equiv \hat{X}' \otimes \hat{Y}' \otimes \hat{Y}'^\dagger$, where $\hat{Y}' \equiv i\hat{Z}\hat{X}'$, calculations yield $\hat{A}|\psi_G\rangle = |\psi_G\rangle$. With $\hat{B} \equiv \hat{Y}'^\dagger \otimes \hat{X}' \otimes \hat{Y}'$ and $\hat{C} \equiv \hat{Y}' \otimes \hat{Y}'^\dagger \otimes \hat{X}'$ also, by symmetry we get $\hat{B}|\psi_G\rangle = |\psi_G\rangle$ and $\hat{C}|\psi_G\rangle = |\psi_G\rangle$. Now $\hat{E} \equiv \hat{A}\hat{B}\hat{C} = \hat{X}' \otimes \hat{Y}'\hat{X}\hat{Y}'^\dagger \otimes \hat{X}'$ and $\hat{D} \equiv \hat{X}' \otimes \hat{X}' \otimes \hat{X}'$ yield the paradox. If values were predefined, the value of $\hat{D}$ and $\hat{E}$ would return the same answer. However, a simple calculation yields $\hat{E}|\psi_G\rangle = |\psi_G\rangle$ (this can be seen directly by applying $\hat{A}$, $\hat{B}$ and $\hat{C}$ sequentially on $|\psi_G\rangle$), while $\hat{D}|\psi_G\rangle = -|\psi_G\rangle$.

The wavefunction now satisfies both the conditions required. The cost that one pays however, is that a simple measurement of $\hat{q}$ and $\hat{p}$ won't suffice. It remains to see precisely which observable one must measure and what must be the analogue of the SG apparatus.

2.3 BM Simulator

Simulation of BM is a two step process. First one must be able to simulate QM, viz. the Schrödinger equation and second, be able to evaluate the position of the particle at each time step.

2.3.1 Design of Numerics

The code was written in Fortran, due to it's efficacy at handling arrays, in conjunction with gnuplot. Runge Kutta 4 was used to solve the differential equations and spline interpolation was used for calculating velocities of the particle. Initial goal was simply to find the trajectories for a single particle, with only one degree of freedom.

2.3.1.1 Simulation of Schrödinger's Equation

To simulate a differential equation, say $\dot{q} = f(q)$, one obvious method is to simply use $q_{n+1} = q_n + f(q_n)\Delta t$, $t_{n+1} = t_n + \Delta t$ where n parametrizes the sequence and Δt, the time step. To get reasonably accurate results, one needs to keep Δt small. This is known as the Euler method and it has errors of $\mathscr{O}(\Delta t^2)$. However, there's another known method, popularly referred to as "RK4", short for Runge Kutta 4, which has errors of $\mathscr{O}(\Delta t^5)$. Assuming again that $\dot{q} = f(q)$, the method claims that $q_{n+1} = q_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$, $t_{n+1} = t_n + h$, where

$$\begin{aligned} k_1 &= f(q_n), \\ k_2 &= f(q_n + hk_1/2), \\ k_3 &= f(q_n + hk_2/2), \\ k_4 &= f(q_n + hk_3), \end{aligned}$$

and $h = \Delta t$. The derivation is tangential to our interest and will have to be skipped. The advantage is that this method can produce accurate results for relatively larger Δt also.

Our purpose is to solve the Schrödinger Equation, viz. $\partial\psi/\partial t = -(\hbar^2/2m)(\partial^2\psi/\partial q^2) + V(q)\psi$. Clearly this is more complicated than the case discussed before, for now we have to solve for ψ, and ψ is a function of both q and t. We begin with uniformly discritizing ψ in the position space, so that ψ is known only at finite points $\{q_n\}$ to start with (and has spacing Δq), while time is discritized as usual, with step Δt. Imagine that we are at the n^{th} step, q_i and t_n are known, and so is $\psi_n(\{q_i\})$ (where the notation means that at the n^{th} step, ψ_n is known at all the points $\{q_i\}$). Given $\psi_n(\{q_i\})$, one can evaluate $\partial^2\psi/\partial q^2 = \psi''$ using the definition of the derivitive, without imposing the $h \to 0$ limit (in the usual notation), to obtain $\psi''(q_i) = [\psi(q_{i+1}) - 2\psi(q_i) + \psi(q_{i-1})]/(\Delta q)^2$. The objective is to find $\psi_{n+1}(q_i)$, which

according to the Euler method can be evaluated as $\psi_{n+1}(q_i) = [-\psi''(q_i) + V(q_i)\psi]\Delta t$, where for illustration, we have set $\hbar^2/2m = 1$. Practically this doesn't suffice and we are forced to use RK4 (or any better alternative). To use RK4, one must be careful and remember that we are evaluating ψ for a given value q_i and therefore in the evaluations of k_i, q_i must be kept fixed, for here ψ is the variable, playing the role of q. Ofcourse, ψ is a complex variable Accordingly, we must have, suppressing q_i to avoid confusion, $\psi_{n+1} = \psi_n + \frac{h}{6}(k_1 + k_2 + k_3 + k_4)$, where

$$\begin{aligned} k_1 &= -\psi_n'' + V\psi_n, \\ k_2 &= -(\psi_n + hk_1/2)'' + V(\psi_n + hk_1/2), \\ k_3 &= -(\psi_n + hk_2/2)'' + V(\psi_n + hk_2/2), \\ k_4 &= -(\psi_n + hk_3)'' + V(\psi_n + hk_3). \end{aligned}$$

where $h = \Delta t$. Let us take a moment to understand how one must go about evaluating this numerically. Since $\psi_n''(q_i)$ and $\psi_n(q_i)$ are known, one can evaluate $k_1(q_i)$ trivially. To evaluate $k_2(q_i)$ however, we would require $k_1(\{q_i\})$ to be known (or at least $k_1(q_{i-1})$, $k_1(q_i)$ and $k_1(q_{i+1})$). Thus, we first evaluate k_1 $\forall q_i$ in the grid. Then to evaluate $k_2(q_i)$, we need $(\psi_n + hk_1/2)''$, evaluated at q_i. Since the finite difference method requires it's argument to be known at q_{i-1}, q_i and q_{i+1}, we see that knowing $k_1(\{q_i\})$ is sufficient to evaluate this step. One similarly evaluates $k_2(\{q_i\})$ to evaluate $k_3(\{q_i\})$ and then $k_4(\{q_i\})$. It maybe added as remark, that practically it is found, that $\Delta t/\Delta x^2 \leq 10^{-6}$ or so for Euler and 10^{-4} or so for RK4, as was also pointed out by [??]. The aforesaid was observed to successfully simulate the Schrödinger equation, for a variety of potentials, details of which follow.

2.3.1.2 Simulation of trajectories

According to Bohmian Mechanics, $\dot{q} = p/m = \nabla S/m = \hbar \mathrm{Im}(\nabla\psi/\psi)$, where $\psi = Re^{iS/\hbar}$. Since from the previous section, we have ψ_n, discretized ψ at each time step, we conclude therefore, that the simplest way to obtain ∇S at an arbitrary position, is to interpolate ψ and subsequently evaluate ∇S. Spline interpolation is quite standard and fairly simple to extend to complex variables. The details will not be discussed here, but the code written [see ??] can be referred to. It maybe mentioned that given $\psi(\{q_i\})$, one can evaluate spline co-efficients, using which, one can easily (computationally cheap) evaluate $\psi(q)$ for practically any value of q. To evolve q, viz. to obtain q_n, again RK4 was used where this time, there no complications. To be able to appreciate better the results however, one needs to plot multiple trajectories simultaneously. The mildly non-trivial aspect of this part, is simply that one needs to start with particles distributed according to $|\psi|^2$ to start with. Attempts were made to find a

library that can do this, but with no luck. A method to convert a unifrom random distribution to an arbitrary distribution was independently arrived at, for a course on Numerical Methods. The idea was to use a uniform random generator, with range $[0,1]$, which for the j^{th} particle, yields $c^{(j)}$, say. The method is that, from the cumulative $\phi(q) = \int_{-\infty}^{q} |\psi(q')|^2 dq'$, one finds $q^{(j)}$ s.t. $\phi(q^{(j)}) = c^{(j)}$, viz. $q^{(j)} = \phi^{-1}(c)$, where the inverse is guaranteed to be single valued, because the cumulative is monotonically increasing. While a formal proof can be presented, we will satisfy ourselves with noting that intuitively, a range of positions with high probability, will cause a larger range of c values to yield those positions, making them more likely to appear, since all values of c are equally likely. With all these ingredients combined, Bohmian trajectories were successfully simulated.

2.3.2 Performance against analytic results

Three known physical situations were tested. (a) $V = 0$; It is known that a Guassian under free evolution, simply expands. (b) $V = \frac{1}{2}\omega^2 q^2$; It is known that a displaced Gaussian, oscillates under a harmonic potential, and so does it's spread (unless a coherent state is chosen). (c) Double slit; By using a time varying potential, s.t.

$$V = \begin{cases} 0 & t \leq t_a \\ V_{\text{dw}} & t_a < t \leq t_b \\ 0 & t_b < t \end{cases}$$

where $V_{\text{dw}}(q)$ represents two wells, simulating the double slit effect, with a single degree of freedom.

In all cases, the results were qualitatively found to match. In case of the double slit, although the wavefunction intereference pattern was distorted, one could see the particles preferentially collecting at the maxima.

2.3.3 Conclusion

BM was successfully simulated for a non-interacting particle, with a single degree of freedom. Implementing RK4, interpolation and random distribution generation were among the non-trivial aspects. Generalization to multiple particles and inclusion of spins, was among the immediate next steps, however due to later theoretical developments, were not carried out.

2.4 Measurements in BM (I)

Measurement of spins in BM was discussed in the previous chapter. Our interest in this chapter, is in measuring q, p and functions there of. We will learn that one can infact construct a universal method that facilitates the measurement of any arbitrary observable. The analysis will clarify two notions which are at high risk of being misconstrued. First, that the position of the particle, when measured, will infact yield q. However, a measurement of momentum of a particle, as we will see, doesn't yield p (in general), the value it had prior to measurement. Infact, it would be inconsistent with QM if it did; for instance, consider $\psi(q) = \left(1/\sqrt{\pi}\right) e^{-q^2}$, for which $S = 0$ and therefore so is $p = \nabla S = 0$. However it is known that upon measurement of p, one obtains a distribution given by the fourier transform of ψ, which, although is centred around the origin, is not a delta function (has non-zero spread). The second notion which is clarified, is that to find the value obtained upon measurement of certain observables, knowledge about the precise position of the measuring particle (the particle used to make the measurement) can play a deciding role.

2.4.1 Observable with discrete spectrum | Hamiltonain Approach

Bohm [??] has discussed in his second paper, how any arbitrary observable maybe measured, in BM. To describe a measurement, we use a measuring particle (mass m_L, say), by getting it to interact with the system for a short duration appropriately, and then measure the final position of the particle, to learn about the value of the observable of interest. Say for instance, we wish to measure the observable $\hat{L}$, then the required interaction Hamiltonian is given by $\hat{H}_{\text{int}} = a\hat{L} \otimes \hat{p}$, where the first operator acts on the system and second (after the tensor product symbol) acts on the measuring particle. a quantifies the interaction strength, and has dimensions of frequency, if L has dimensions of length. Let us assume that the system in the state $|\psi\rangle$. Then it is known that one can express $|\psi\rangle = \sum_l \langle l|\psi\rangle |l\rangle$, where $|l\rangle$ are the eigenstates of $\hat{L}$ with eigenvalue l (we have assumed non-degenercy for simplicity, but its removal doesn't cause any significant difficulty). $|\Psi_S(t)\rangle = U(t)|\psi\rangle \otimes |\varphi\rangle$, where

$$U(t) = e^{-\frac{i}{\hbar}\left[-\hbar^2\frac{\nabla_1^2}{2m} - \hbar^2\frac{\nabla_2^2}{2m_L} + \hat{H}_{\text{int}}\right]t}$$

and $|\varphi\rangle$ is the state of the particle, given by a Guassian centred at the origin, $\varphi(q) = (1/\sqrt{2\pi}\sigma)e^{-q^2/2\sigma^2}$. If t is very small, and a very large, s.t. $at = \lambda$ is a finite number, then one can neglect the free

evolution, $\nabla^2 t$ terms compared to $\hat{H}_{\text{int}} t$, to obtain

$$\begin{aligned} |\Psi_S(t)\rangle &= e^{-\frac{i}{\hbar} a\hat{L}\otimes\hat{p}t} |\Psi_S\rangle \\ &= \sum_l \langle l|\psi\rangle |l\rangle \otimes e^{-\frac{i}{\hbar}\lambda l \hat{p}} |\varphi\rangle \\ &= \sum_l \langle l|\psi\rangle |l\rangle \otimes |\varphi_{\lambda l}\rangle , \end{aligned}$$

where $\left|\varphi_{q_0}\right\rangle = \int dq \varphi(q-q_0)|q\rangle$. This interaction, effectively entangles the measuring particle, with the possible 'outcomes', eigenstates of the observable $\hat{L}$ of interest. If $\sigma \ll \lambda l$, then according to QM itself, a position measurement of the measuring particle, would correspond, in a one-to-one way, to the eigenstate/eigenvalue of $\hat{L}$, to which the system will collapse. In the context of BM, after the interaction, the measuring particle would be guided into one of these eigenstates, and a position measurement would yield the same. We glossed over various details in making the last statement, which will be delineated shortly, through some illustrations.

2.4.1.1 Generalization to observables with continuum spectrum

Since we are interested in observing phase space variables, the aforesaid must be generalized carefully to the continuous spectra regime. (TODO: write this from the notes)

2.4.1.2 Consistency check; measurement of position

The position of a particle, as is claimed by BM, remains the same upon observing. Let us verify this statement, by applying the aforesaid formalism. For simplicity, let us assume that the state of the system is given as $\sqrt{2}|\psi\rangle = \int dq\,[\delta(q-q_0)+\delta(q+q_0)]|q\rangle$, which is to say, the system is in a superposition of being at q_0 and $-q_0$. Let us assume that the particle is initially given to be in $q_1 = q_0$. Now our measurement formalism must be consistent with this, viz. the answer must not depend on the initial position of the measuring particle. Initially then, the state is $|\Psi_S\rangle = |\psi\rangle \otimes |\varphi\rangle$, while after interaction, the state becomes $\sqrt{2}|\Psi_S(t)\rangle = |q_0\rangle \otimes \left|\varphi_{\lambda q_0}\right\rangle + |-q_0\rangle \otimes \left|\varphi_{-\lambda q_0}\right\rangle$. To understand this clearly, one must plot $|\Psi_S(q_1,q_2)|^2 = |\langle q_1,q_2|\Psi_S\rangle|^2$, where q_1 is (the position eigenket) for the system and q_2 for the measuring particle. Before the interaction [see Fig. ??], say the measuring particle was at $q_2 \in (-\sigma,\sigma)$. After the interaction [see Fig. ??], the measuring particle will move to $q_2 \in (\lambda q_0 - \sigma, \lambda q_0 + \sigma)$, which follows simply by noting that the velocities are given by probability current. A measurement of the particle's position would yield $q_2 \approx \lambda q_0$, which we know from our analysis, corresponds to $q_1 = q_0$, which is infact consistent with our initial conditions. It must be emphasized that the position of the measuring particle, plays no essential

role in deciding where it would show up. The same analysis can be repeated for $q_1 = -q_0$ and in this case, from Fig. ?? it would follow that after the interaction, $q_2 \approx -\lambda q_0$. Consistency of this method is therefore established. Ofcourse, one can relax the idealized δ functions to obtain more realistic functions, but the conclusion remains invariant.

2.4.1.3 Consistency check; measurement of momentum

To resolve the issue pointed out in the beginning of this section, viz. the observed value of p being inconsistent with QM, we begin with writing the state of the system as $|\psi\rangle = \int dq \left(1/\sqrt{\pi}\right) e^{-q^2} |q\rangle$, so that $p = \nabla S = 0$. We now wish to find, how the spread in the observed value of p appears, from the aforesaid measurement formalism. Note that

$$\begin{aligned} |\psi\rangle &= \int dp\,|p\rangle \int dq \langle p|q\rangle \overbrace{\left(1/\sqrt{2\pi\hbar}\right) e^{-iqp/\hbar}}\psi(q) \\ &= \int dp \tilde{\psi}(p)\,|p\rangle\,, \end{aligned}$$

where $\tilde{\psi}(p)$ is $\psi(q)$ fourier transformed. The combined state before measurement is given by $|\Psi_S\rangle = |\psi\rangle \otimes |\varphi\rangle$, while after the measurement, $|\Psi(t)\rangle = \int dp \tilde{\psi}(p)\,|p\rangle \otimes \left|\varphi_{\lambda p}\right\rangle$. The aforesaid expression entails that, assuming $\sigma \to 0$, the probability for the measuring particle, to end up at λp, is given by $|\tilde{\psi}(p)|^2$, which is consistent with QM. To obtain this result more carefully from Bohmian trajectories, one must plot $|\Psi_S(q_1, q_2)|^2$ as was done in the previous section, but we will satisfy ourselves here, with noting that the measurement process has explicitly imparted momentum on the system, as is clear after 'collapsing' the wavefunction.

2.4.2 Classical limit of measurements

Although we haven't yet discussed the classical limit of BM, as an aside, it maybe demonstrated that this indeed follows elegantly. If we write $\psi = Re^{iS/\hbar}$ and substitute this in the Schrödinger Equation, and separate the real and imaginary parts, we obtain

$$\begin{aligned} \frac{\partial R}{\partial t} &= -\frac{1}{2m}\left[R\nabla^2 S + 2\nabla R.\nabla S\right], \\ \frac{\partial S}{\partial t} &= -\left[\frac{\nabla S^2}{2m} + V - \frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}\right]. \end{aligned}$$

The reader would've recognized that the second equation is essentially the Hamilton Jacobi equation, with an extra term. Infact, if one writes $P = R^2$, then we obtain

$$\begin{aligned} \frac{\partial P}{\partial t} + \nabla.\left(P\frac{\nabla S}{m}\right) &= 0 \\ \frac{\partial S}{\partial t} + \frac{\nabla S^2}{2m} + V - \frac{\hbar^2}{4m}\left[\frac{\nabla^2 P}{P} - \frac{1}{2}\frac{\nabla P^2}{P^2}\right] &= 0 \end{aligned}$$

which makes the first equation effectively a continuity equation for probabilities, if one relates $\nabla S = p$ (momentum in BM), as is done in the Hamilton Jacobi framework. Effectively then, BM elegantly yields the classical limit, for if we neglect the $\hbar$ term, then we have particles obeying Newton's laws.

Our motivation was to study how the measured value of momentum turns out to be $p = \nabla S$ in the classical limit. The larger goal, was to use this elegant framework, to understand how measurement of functions of q, p, such as the energy for example, become equivalent to measuring q and p and plugging their values into the function. To be precise, consider the example of an energy eigenstate in a harmonic potential. For this state, neither q nor p are precisely defined (in the language of QM), however, the energy is sharp. We started with analysing the classical limit of the measurement process, just discussed. That requires the potential V, to be a function of both q and p, whereas for the aforesaid derivation, only position dependent V was used. Correcting for this, create various issues. First of all, this interaction adds a term to the continuity equation,

$$\frac{\partial P}{\partial t} + \nabla_1\left(P\frac{\nabla_1 S}{m}\right) + \nabla_2\left(P\frac{\nabla_2 S}{m}\right) + \overbrace{2aR\left[\nabla_1 R\nabla_2 S + R\nabla_1\nabla_2 S + \nabla_2 R\nabla_1 S\right]}^{\text{Extra Term! | How will } P \text{ satisfy the continuity equation?}} = 0.$$

The interaction also adds a "quantum potential" in addition to the expected potential, which doesn't seem to disappear in the large mass limit,

$$\frac{\partial S}{\partial t} + \frac{(\nabla_1 S)^2}{2m} + \frac{(\nabla_2 S)^2}{2m} + \overbrace{a\nabla_1 S\nabla_2 S}^{\text{expected part}} - \overbrace{a\hbar^2\frac{\nabla_1\nabla_2 R}{R}}^{\text{quantum part}} - \underbrace{\frac{\hbar}{2mR}(\nabla_1^2 R + \nabla_2^2 R)}_{\text{usual quantum potential}} = 0.$$

One can see that with $\hbar \to 0$, the quantum potential and the quantum part of the interaction, both disappear. However, the continuity equation is not recovered still. This part was not explored further, but can perhaps be studied independent of our current target.

2.4.3 Conclusion

From the optimized GHZ construction, it was necessary to arrive at a theoretical construction, that would allow calculation of measurement outcomes associated with arbitrary operators composed using $\hat{q}$, $\hat{p}$. This was achieved, however it was found that obtaining these results is not trivial in most cases of interest. The surprising result, that measurement of $\hat{p}$ may not yield $p = \nabla S$, while a measurement of $\hat{q}$ yields q, was derived and clarified. This takes us a step closer to understanding how BM can explain the optimized (or even the phase space) GHZ test, by allowing an in principle simulation.

2.5 Roy Singh Theory

Roy Singh proposed a theory that was a causal completion of QM, that treated position and momentum symmetrically, viz. their joint probability distribution was such that when the momentum is integrated out, the result agrees with quantum mechanical predictions for position *and* when positions are integrated out, they agree with the quantum mechanical momentum distribution. This was rather interesting for, BM fails to do the latter, and requires more analysis to resolve (as was discussed in the previous section). This joint probaiblity distribution, given by RS, could indeed be thought of as describing the phase space of real particles, for it was positive everywhere, unlike the Wigner distribution. The cost however was that evaluation of arbitrary functions of q, p was not as simple as integrating it over the joint distribution. While a promising framework and interesting in its own right, in the light of the optimized GHZ test, it wasn't particularly helpful and not pursued beyond a preliminary reading stage. (TODO: Add formalism if possible)

2.6 Concluding Remarks

In summary then, we have looked at a version of the GHZ test, the phase space extension, that effectively says that $\hat{q}, \hat{p}$ can't have pre-defined values. Now, even though BM has pre-defined values for q, p in principle, upon measurement, we learnt that the value of p can change, and may very well depend on the initial position of the measuring particle. To be able to test this quantitatively, we had to (a) optimize the phase space GHZ test, (b) learn more about measurements and (c) write an appropriate BM simulator. Progress was made on each front and the results discussed. However, the crucial point is that even in BM, the measurements correspond to operators. It entails therefore that measuring $ABC \neq D$ but infact $ABC = -D$, (where A, B, C, D are as defined for the spin GHZ, and the argument essentially holds even

for the phase-space (optimized) GHZ) even from the point of view of BM, a deterministic theory. This hinges essentially on whether we can commute the operators that were used to construct A, B and C to start with, which even in BM we can't. Stated another way, we learn that $YXY \neq Y^2X = X$, but $YXY = -Y^2X = -X$, even in BM, where these operators have pre-defined values (given all initial conditions), simply because X and Y are operators and they anti-commute. So even though numerically we haven't quite proven that BM will be consistent, qualitative arguments are already convincing enough that BM will infact turn out to be consistent with QM and what remains is a matter of detail.

Chapter 3

Contextuality and Bohmian Mechanics

3.1 Rationale

We saw in the last chapter that BM is able to explain the GHZ test, essentially because the paradox arises if we try to multiply values assigned to operators that don't commute, to obtain the value assigned to a product of such operators. In BM, since the operator must be provided to obtain it's value, no contradiction arises, even though the results are deterministic, given all initial conditions. Effectively then, the flaw in our GHZ reasoning, was that we were commuting operators that didn't commute, by assuming they were numbers. Thus, we were only able to rule out theories that treated operators like numbers and we find that BM is not one of them. However, as we have seen in the Peres Mermin (PM) test, the operators whose values are multiplied, do infact commute, implying that the PM test is exonerated from that objection. And we also know that BM, atleast on the face of it, is non-contextual, viz. the value obtained upon the measurement of an operator, depends only on the state (+ hidden variables) and the operator. It is therefore not clear how BM is contextual, which is a notion we're forced to accept from the PM test. Further, since BM is claimed to reproduce all results of QM, does it entail that a non-contextual theory can explain the PM test?

We also found in the previous chapter, that in BM, there's no fundamental difference that arises between the phase space variables and spins, for as we saw, the value of momentum observed, depends on the systematics of the measurement process, just as the outcome of a spin measurement using an SG setup (as was discussed in the first chapter). Therefore in what follows, attempts will be made at tackling the problem from whichever method appears more accessible.

3.2 Intensifying Contextuality

We have already seen that GHZ test is a test of determinism, of whether predefined values can explain QM, while the PM test is that of contextuality, which says that predefined non-contextual values, a more subtle remark, can't explain QM. We also learnt how one can extend the GHZ test to continuous variables. In this section, we will make both these schemes effectively equally powerful; make GHZ into a test of contextuality and extend the PM test to continuous variables.

3.2.1 GHZ escalated to Contextuality

Here, we present a successful attempt at converting the GHZ test, into a full fledged test of contextuality. Recall from the first chapter, $\hat{A} \equiv \sigma_x \otimes \sigma_y \otimes \sigma_y$, $\hat{B} \equiv \sigma_y \otimes \sigma_x \otimes \sigma_y$, $\hat{C} \equiv \sigma_y \otimes \sigma_y \otimes \sigma_x$ and $\hat{D} = \sigma_x \otimes \sigma_x \otimes \sigma_x$, and $\hat{A}\hat{B}\hat{C} = -\hat{D}$. Consider now, in addition, the following operators,

$$\hat{H}_{ij} \doteq \begin{bmatrix} \sigma_x \otimes \mathbb{I} \otimes \mathbb{I}^{(a)} & \mathbb{I} \otimes \sigma_y \otimes \mathbb{I}^{(2)} & \mathbb{I} \otimes \mathbb{I} \otimes \sigma_y^{(3)} \\ \sigma_y \otimes \mathbb{I} \otimes \mathbb{I}^{(1)} & \mathbb{I} \otimes \sigma_x \otimes \mathbb{I}^{(b)} & \mathbb{I} \otimes \mathbb{I} \otimes \sigma_y^{(3)} \\ \sigma_y \otimes \mathbb{I} \otimes \mathbb{I}^{(1)} & \mathbb{I} \otimes \sigma_y \otimes \mathbb{I}^{(2)} & \mathbb{I} \otimes \mathbb{I} \otimes \sigma_x^{(c)} \\ \sigma_x \otimes \mathbb{I} \otimes \mathbb{I}^{(a)} & \mathbb{I} \otimes \sigma_x \otimes \mathbb{I}^{(b)} & \mathbb{I} \otimes \mathbb{I} \otimes \sigma_x^{(c)} \end{bmatrix}$$

and note that (a) the product of these operators along each row, yields A B C and D respectively, (b) operators along any row, commute and (c) the superscript labels identify operators that are repeated.

We will now try to assign values to $\hat{H}_{ij}$ and only multiply values assigned to commuting observables, to obtain the value of $\hat{A}, \hat{B}, \hat{C}$ and $\hat{D}$. We know (see Section [??]) that according to QM, for the GHZ state, $|\chi_G\rangle$, $\hat{A}$, $\hat{B}$ and $\hat{C}$ must be assigned the value $+1$, while $\hat{D}$ must be assigned -1. We demand that our assignment corresponds to this state. Let us start with assigning values to the last row. To obtain a -1 corresponding to D, we must have either a $1, 1, -1$ (or permutations) or $(-1, -1, -1)$. Let us assume that the former is true. At this stage then, we'd have

$$H_{ij} \doteq \begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \\ 1 & 1 & -1 \end{bmatrix}.$$

Now to obtain a $+1$ in the first row (for A), we must have $(1, \pm 1, \pm 1)$. Similarly we get, for

the second row, $(\pm 1, 1, \pm 1)$. The assignment then becomes

$$H_{ij} \doteq \begin{bmatrix} 1 & \pm 1 & \pm 1 \\ \pm 1 & 1 & \pm 1 \\ \pm 1 & \pm 1 & -1 \\ 1 & 1 & -1 \end{bmatrix},$$

however, we have no freedom left and are forced to assign -1 to the third row, while we were required to have it equal $+1$. A little reflection reveals that the $(-1,-1,-1)$ case for the last row, won't resolve the issue. Thus we conclude that the assignment must be contextual (atleast the one corresponding to $|\chi_G\rangle$). In this scenario, the non-commuting observables argument, deduced in the previous chapter fails, for values of only commuting observables were multiplied, even though, effectively it is still the same GHZ test.

3.2.2 Peres Mermin escalated to Continuous Variables

The Peres Mermin situation can be, quite simply, extended to continuous variables, which maybe of practical interest, since performing SG type setup in the simulation requires implementation of spins and appropriate magnetic field effects, and that can be surpassed with this construction. To this end, the following extension was worked out (but this result was already known). For any unitary operators X and Y, s.t. $\{X,Y\} = 0$ and so is $\{X,Y^\dagger\} = 0$, if we define $Z = iYX$, then

$$\begin{array}{ccc} X \otimes \mathbb{I} & \mathbb{I} \otimes X & X^\dagger \otimes X^\dagger \\ \mathbb{I} \otimes Y & Y \otimes \mathbb{I} & Y^\dagger \otimes Y^\dagger \\ X^\dagger \otimes Y^\dagger & Y^\dagger \otimes X^\dagger & Z \otimes Z \end{array}$$

would yield the PM situation. To check this, note that the product along any row, is $\mathbb{I}$ and it is also so for each column, except for the third, which is $-\mathbb{I}$, precisely the same as the PM situation. The difference however, will be that the corresponding hidden variables, will have to be unimodular complex (viz. $XX^* = 1$) and the values of the operators, must be deduced from their hermitian counterparts. One such choice of X and Y was pointed out in the previous chapter; $\hat{X} \equiv e^{i\sqrt{\pi}\hat{q}/\hbar L}$ and $\hat{Y} \equiv e^{i\sqrt{\pi}\hat{p}L/\hbar}$. Note that in this case, optimization is not required, for this test is state independent (should work with any choice of states).

3.3 Measurements in BM (II)

Measurements have already been discussed at length in the previous chapters, however, we are now particularly interested in measuring spins, especially when there is entanglement. We

will see how a measurement made using the Stern Gerlach like apparatus can yield results very different from those obtained by the Hamiltonian approach. Further, we will see how the former can get difficult to evaluate, quite quickly.

3.3.1 Stern Gerlach based measurements

Since we have already discussed how to analyse a Stern Gerlach based measurement for a single particle, let us try to explore how to analyse this for $|\chi\rangle = |00\rangle + |11\rangle/\sqrt{2}$, where for instance $|00\rangle = |0\rangle_1 \otimes |0\rangle_2$ (1 and 2 label the particle), and $\hat{\sigma}_z \otimes \hat{\sigma}_z$ is the observable of interest. Borrowing the notation and the analysis from Section 1.4, we have $\sqrt{2}\,|\Psi_S\rangle = |\psi\psi\rangle \otimes [|00\rangle + |11\rangle]$ initially and after the interaction, $\sqrt{2}\,|\Psi_S(t)\rangle = |\psi_{at}, \psi_{at}\rangle \otimes |00\rangle + |\psi_{-at}, \psi_{-at}\rangle \otimes |11\rangle$, where $|\psi_{at}, \psi_{at}\rangle = |\psi_{at}\rangle_1 \otimes |\psi_{at}\rangle_2$ for example. Now, we must predict the outcome, given the initial positions of the particles. We then plot $P(q_1, q_2) = \langle \Psi_S | q_1, q_2 \rangle \langle q_1, q_2 | \Psi_S \rangle$, where q_1 and q_2 represent the z coordinate of the particles, and effectively we are summing over the spin degree of freedom, asking for the probability of finding the particles near q_1 and q_2. Initially, $P = |\langle q_1, q_2 | \psi\psi \rangle|^2$ which would yield circles in a contour plot (showing say 70% probability enclosed within), see Fig. [??]. After the interaction, $2P(q_1, q_2) = |\langle q_1, q_2 | \psi_{at}, \psi_{at} \rangle|^2 + |\langle q_1, q_2 | \psi_{-at}, \psi_{-at} \rangle|^2$, which would yield two circles in the countour plot, one centred at (at, at) and the other centred at $(-at, -at)$, see Fig. [??]. From the symmetry of the problem, and from the fact that velocities are single valued, we can deduce that if the particle positions were represented by a point inside area I (see Fig. [??]), then the result would be a spin up, viz. $|0\rangle$ for both, while for the other case, the result would be spin down for both. Note that there's no initial condition for which the results are anti-correlated, consistent with $|\chi\rangle$. Although it is obvious, it is instructive to see how this would work for $|\chi\rangle = |00\rangle$ for instance. In this case also, the initial plot for P is identical to that shown in Fig. [??]. However, since after the interaction, $|\Psi_S(t)\rangle = |\psi_{at}, \psi_{at}\rangle \otimes |00\rangle$, we see that $P = |\langle q_1, q_2 | \psi_{at}, \psi_{at} \rangle|^2$, which is just a Gaussian centred at (at, at). Now, regardless of the initial conditions of the particles, after the interaction, they will be carried to a Gaussian at (at, at) by the probability current (which is effectively their velocity). Thus, the result is always spin up, for both particles, consistent with QM.

It is straight forward to extend this analysis to the GHZ situation, with three particles. However, in that case, determining which area (in phase space, viz. the set of initial positions) corresponds to which outcome, becomes non-trivial, for certain operators. The difficulty arises from the increasing possible final states.

3.3.2 Hamiltonian Approach

We have already used the Hamiltonian approach in Section 2.4 and so extending it to spins will not be surprising. However, as we will see, it incredibly simplifies the measurement process.

3.3.2.1 Single Spin

Let the spin state of our particle, be $\sqrt{2}\,|\chi\rangle = |0\rangle + |1\rangle$ and let the spatial state of the measuring particle be $|\psi\rangle$. Neglecting the spatial wavefunction of the particle of interest, we can write the state of the entire system as $|\Psi_S\rangle = |\psi\rangle \otimes |\chi\rangle$. From Section 2.4, we know that the required Hamiltonian is $\hat{H} = a\hat{\sigma}_z \otimes \hat{p}$, given that we wish to measure $\hat{\sigma}_z$. Consequently, after the interaction, we get $\sqrt{2}\,|\Psi_S(t)\rangle = |\psi_{at}\rangle \otimes |0\rangle + |\psi_{-at}\rangle \otimes |1\rangle$. This has become mathematically identical to the SG based measurement discussed in Section 1.4. The results therefore, follow from our older discussions. However, to stress the difference, note that in the SG case, the position of the particle of interest itself determined whether the measurement would yield spin up or spin down. Here, position of the particle of interest, plays no role in the measurement scheme, infact, the position of the measuring particle (which in some sense represents the apparatus) determines the outcome. Thus, even though in both cases, the result can be predicted, it is possible to have initial conditions such that the results of these two different measurement schemes, don't match. In the following, we will see how this scheme, relatively simplifies the measurement process.

3.3.2.2 Entangled Spin

Let the spin state of the particles be $\sqrt{2}\,|\chi\rangle = |00\rangle + |11\rangle$ and let the position of the measuring particle be $|\psi\rangle$. Neglecting again, the spatial wavefunction of the two particles, we have, for the entire system, $|\Psi_S\rangle = |\psi\rangle \otimes |\chi\rangle$. If we wish to measure $\hat{\sigma}_z \otimes \hat{\sigma}_z$, which must yield $+1$, we must use $\hat{H} = a(\hat{\sigma}_z \otimes \hat{\sigma}_z) \otimes \hat{p}$. This yields $|\Psi_S(t)\rangle = |\psi_{at}\rangle \otimes |\chi\rangle$, which from the aforesaid arguments, should be clear, entails that for all initial positions of the measuring particle, we would get $+1$. Since the result is $+1$, we conclude further that the spins must be correlated, much more quickly than we could using the SG setup.

3.3.2.3 GHZ Entangled Spin

Let the spin state, now be $\sqrt{2}\,|\chi\rangle = |000\rangle - |111\rangle$, the case we couldn't analyse in a simple way earlier. Assume that we wish to measure $\hat{A} = \hat{\sigma}_x \otimes \hat{\sigma}_y \otimes \hat{\sigma}_y$. Since $\hat{A}\,|\chi\rangle = |\chi\rangle$, it follows that after the interaction, we would have $|\Psi_S(t)\rangle = |\psi_{at}\rangle \otimes |\chi\rangle$, that'll yield $+1$ with certainty. Obtaining this conclusion from the SG formalism would've been much harder. Let

us attempt something non-trivial; let us measure $\hat{\sigma}_x \otimes \mathbb{I} \otimes \mathbb{I}$. To proceed, we re-write the spin state as $\sqrt{2}|\chi\rangle = |+\rangle(|00\rangle - |11\rangle) + |-\rangle(|00\rangle + |11\rangle)$ and now observe that the state of the complete system, after the interaction, would be $\sqrt{2}|\Psi_S(t)\rangle = |\psi_{at}\rangle \otimes |+\rangle(|00\rangle - |11\rangle) + |\psi_{-at}\rangle \otimes |-\rangle(|00\rangle + |11\rangle)$. Now if the measuring particle was initially at $q > 0$, we would get a $+1$ value and the resultant state would be $|+\rangle(|00\rangle - |11\rangle)$ and similarly for $q < 0$, -1 and $|-\rangle(|00\rangle + |11\rangle)$.

3.4 Peres Mermin Revisited

With the tools for measurement sharpened, we are now in position to explicitly assign values to the Peres Mermin situation and find out precisely how contextuality can be reconciled with BM. However, before proceeding with that, we first try to carefully list the restrictions made on the hidden variable theory, which are sufficient to arrive at a contradiction with QM.

3.4.1 Assumptions

According to QM, if one has a set of compatible operators, say $\hat{A}, \hat{B}, \hat{C}$ (some arbitrary operators), then one can construct eigenkets$|a, b, c\rangle$. Now if the system is prepared in such a state, then the value obtained by measuring $\hat{A}\hat{B}\hat{C}$ is the same as that obtained by measuring $\hat{A}, \hat{B}$ and $\hat{C}$ separately, in any order, and multiplying the result. We can arrive at a contradiction, if we make the following three assumptions about the HV model. (1) Non-contextual assignment: The value assigned to any operator, depends only on the operator and the state of the system (including hidden variables). (2) Value assigned to the product of commuting operators must be the product of values assigned to the commuting operators themselves. (3) Measurement doesn't affect the remaining assignments.

It is already clear that relaxing requirements (2) and (3) might make the assignment consistent with QM, voiding the necessity of concluding the necessity of having a contextual assignment.

3.4.2 BM explanation of Peres Mermin

Let us explicitly assign values to the PM situation to explore precisely which of the aforesaid assumptions goes wrong. For simplicity, let us assume that our state to start with, is $|\chi\rangle = |00\rangle$ and $q > 0$ for the measuring particle, initially. For convenience, we have re-written the relevant

operators

$$\hat{A}_{ij} \doteq \begin{bmatrix} \hat{\mathbb{I}}\otimes\hat{\sigma}_x & \hat{\sigma}_x\otimes\hat{\mathbb{I}} & \hat{\sigma}_x\otimes\hat{\sigma}_x \\ \hat{\sigma}_y\otimes\hat{\mathbb{I}} & \hat{\mathbb{I}}\otimes\hat{\sigma}_y & \hat{\sigma}_y\otimes\hat{\sigma}_y \\ \hat{\sigma}_y\otimes\hat{\sigma}_x & \hat{\sigma}_x\otimes\hat{\sigma}_y & \hat{\sigma}_z\otimes\hat{\sigma}_z \end{bmatrix}.$$

From our earlier analysis, we know that operators for which $|\chi\rangle$ is an eigenvector, they will always yield the corresponding eigenvalue. In the present case, only $\hat{\sigma}_z\otimes\hat{\sigma}_z$ has $|\chi\rangle$ as an eigenket. Thus we assign it $+1$ (the eigenvalue). We will not present the details of calculations that yield the following predictions. They can however be easily motivated, by noting that each operator, can yield only a ±1 value, since $\hat{A}_{ij}^2 = \hat{\mathbb{I}}\otimes\hat{\mathbb{I}}$. By symmetry of the operators and the state, it follows that (one can check explicitly) each of these operators (for which $|\chi\rangle$ is not an eigenket), yield ±1 with equal probability. Thus, the eigenket of these operators that corresponds to the $+1$ value, is the one that will correspond to $|\psi_{at}\rangle$. Since $q > 0$ by assumption, it follows that all these operators would yield $+1$, viz. the assignment will be

$$\begin{bmatrix} +1 & +1 & +1 \\ +1 & +1 & +1 \\ +1 & +1 & +1 \end{bmatrix}.$$

It is obvious that this assignment, if we assume (2), will fail to be consistent with QM. However, since $\hat{R}_i = \mathbb{I}$ and $\hat{C}_j = \mathbb{I}\,(j \neq 3)$, $\hat{C}_3 = -\mathbb{I}$, $(\forall\, i, j)$ have $|\chi\rangle$ as an eigenket (infact all possible states are eigenkets), according to BM, the value assigned will be $+1$ to all except C_3, which will be assigned -1. We learn therefore that in BM, (2) certainly doesn't hold. Also, since the wavefunction collapses after a measurement, we know that the assignment must change in general, thus (3) also doesn't hold in BM. Finally, given an operator, BM can predict the value one would get upon measuring it. Thus we see, that (1) does infact hold in BM.

3.5 Conclusion

It was realized that the GHZ determinism test is not satisfactorily explained by the non-commuting observables being treated as numbers. An equivalence between the GHZ and the PM test was established, in terms of the conclusions. Consequently, tools required to find an explicit assignment for the PM situation were sharpened and from the assignment, it was found that a non-contextual theory doesn't contradict QM, viz. contextuality is not necessary.

Chapter 4

An Alternative to Contextuality

4.1 Rationale

4.2 Multiplicativity

4.3 Simple Models

4.3.1 Contextual, Memory Model

4.3.2 Non-Contextual, Toy Model

4.4 Generic Models

4.4.1 Generalized Toy Model

4.4.2 C-ingle Theory

4.5 Conclusion

Chapter 5

Epilogue

5.1 The Verdict: Contextuality vs. Bohmian Mechanics

5.2 Summary of Results

5.3 Conclusion

5.4 Digressions

5.4.1 Attempt at superluminal communication with BM

5.4.2 Identical Particles, alternative approach

References

Appendix A

Bohmian Mechanics, from Bohm's papers

The usual interpretation of quantum mechanics (QM) is self consistent. It assumes however that the most complete possible specification of an individual system is in terms of ψ that yields only probabilistic results. To investigate if this assumption is accurate, it is reasonable to expect an experimental test to exist. However, as will be shown, no experiment can test this interpretation. The other alternative then, is to attempt constructing a theory, which has extra variables which in principle completely specify the system but in practice get averaged over. If such a theory existed then it is obvious that the assumption false. In his paper, Bohm describes such a theory. This theory yields the same results as QM in all physical situations studied and is shown to be broader conceptually. The mathematics used is shown to be more general. Bohm's idea was that while QM fails at high energies, his theory might work. In retrospection though, one realizes how both BM and QM suffer from the same problem which requires quantum field theory (QFT) to remedy.

A.1 Introduction

The usual interpretation assumes that the wavefunction is the most complete description. This description was criticized by Einstein for he believed that there must exist a better, precise theory that describes nature as opposed to QM that claims nature is fundamentally probabilistic. Einstein's views have themselves been criticised to be irrelevant because the usual QM interpretation is in excellent agreement with experiments.

Appendix B

Bohmian Mechanics, from Holland

It is worthwhile spending some time trying reformulate classical mechanics (CM) into a form that is closest to quantum mechanics. Although this is not necessary for understanding BM (and may very well be skipped), doing so brings with it it's own perspective and excitement.

B.1 The Hamilton Jacobi Theory

B.1.1 Introduction

Let's start with a Lagrangian, $L = L(q,\dot{q},t)$, where $p_i := \partial L/\partial \dot{q}_i$, $H := p_i\dot{q}_i - L$, q_i represents the position of the ith particle and $q,\dot{q}$ refer to the set $\{q_1,q_2,\dots\}$, $\{\dot{q}_1,\dot{q}_2,\dots\}$ respectivley. If we imagine a path that the particle takes from a point q_0,t_0 to q,t, then we can define a quantity $I(q,t;q_0,t_0) := \int_{q_0,t_0}^{q,t} L(q,\dot{q},t)dt$ where the integral is taken along the aforesaid path. We know from experiments that the path the particle actually takes, given L, is one which extremizes I. This entails that, just like at the extremum value of the function, it's derivative vanishes, the first order variation of I for a path that extremizes I must vanish; viz. $\delta I = 0$ for any extremal path. As the reader would recall, after some calculation, this yields

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0.$$

This can be expressed in terms of H instead of L by substituting for L in the definition of I and changing the independent variables from $(q,\dot{q})$ to (q,p) & varying them, where p refers to $\{p_1,p_2,\dots\}$. The result is the familiar Hamilton equations (this again requires some calculation and may not always be possible)

$$\dot{q}_i = \frac{\partial H}{\partial p_i},\ \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

Defn: Canoncial Transformation: $q,p \to Q = Q(q,p,t), P = P(q,p,t)$ is a a canonical transformation if Q, P preserve the form of the Hamilton Equations, viz. if $K = K(Q,P,t)$ is the new H, then we must have $\dot{Q}_i = \frac{\partial K}{\partial P_i}, \dot{P}_i = -\frac{\partial K}{\partial Q_i}$ which could've come from $I = \int P_i dQ_i - K dt$.

It is known that if two Ls differ by $\frac{df}{dt}$, where f [check: where $f = f(q,\dot{q},t)$] is some function, then they yield identical equations of motion. [) This follows at once by recalling that $I_1 = \int_{q_0,t}^{q,t} L_1 dt$ and noting that $I_2 = \int_{q_0,t}^{q,t} L_1 dt + F - F_0$ entails that the variation $\delta I_2 = \delta I_1 + \cancelto{0}{\delta F} - \cancelto{0}{\delta F_0} = \delta I_1$(]. Thus if we write the Lagrangian in terms of the Hamiltonian and impose

$$P_i dQ_i - K(Q,P)dt + \frac{dF}{dt} = p_i dq_i - H(q,p)dt,$$

then we're guaranteed that the equation of motion (EOM) will be identical and hence describe the same physical system. Assuming a specific functional form, viz. $F = F(q,Q,t)$, one can evaluate

$$\begin{aligned} p_i &= \frac{\partial F}{\partial q_i}, \\ P_i &= -\frac{\partial F}{\partial Q_i}, \\ K &= H + \frac{\partial F}{\partial t}. \end{aligned}$$

[) This follows from the restatement $\left(P_i + \frac{\partial F}{\partial q_i}\right)\dot{Q}_i - \left(p_i - \frac{\partial F}{\partial q_i}\right)\dot{q}_i + \left(K - H - \frac{\partial F}{\partial t}\right) = 0$ and by choosing q, Q to be the independent variables. (] Note that the first equation will yield Q in terms of p and q. The following equation yields P in terms of Q and q. The last equation yields K, since Q and P are known (it's not obvious though that writing $K = K(Q,P)$ will be straight forward).

Thus, given $F(q,Q,t)$ we can find the canonical transformation $Q = Q(q,p,t), P = P(q,p,t)$ and even $K = K(q,p,t)$. Defn: (by example) F is a generating function.

Next let us consider a case where the transformed variables are almost the same as the original set, viz.

$$\begin{aligned} Q_i &= q_i + \delta q_i, \\ P_i &= p_i + \delta p_i. \end{aligned}$$

The claim is that $\delta q_i = \varepsilon \partial G/\partial p_i$ and $\delta p_i = -\varepsilon \partial G/\partial q_i$, where $G = G(q,p,t)$ is the generating function and ε is small. [) This is plaussible by simply taking a generating function which is a function of (q,p,t) instead of (q,Q,t) as was F (] Note that for $\varepsilon = dt$, $G = H$, it follows

$\partial q = dt\partial H/\partial p_i = dt\dot{q}_i = dq_i$ and similarly $\delta p_i = dp_i$. The central idea is that infinitesimal canonical transformations $\leftrightarrow$ infinitesimal motion. Successive infinitesimal canonical transformations yield a finite canonical transformation. Successive infinitesimal motion describes observed motion. Consequently, any observed classical motion can be described by an appropriate canonical transformation. This can be used to find the motion itself, in fact.

One tries to find a canonical transformation corresponding to $q = q(q_0, p_0, t)$, $p = p(q_0, p_0, t)$ which can be inverted to yield q_0, p_0(at t_0) given q, p, t. [] Thus, all points along the trajectory, viz. the set (q, p, t), all map to the same point (q_0, p_0, t_0). [] The mapped point, must not evolve with time. Thus, we demand (or require) that $K = 0$ (basically must be constant, can be $\neq 0$) so that $\dot{Q} = 0, \dot{P} = 0$. Conventionally, the associated generating function is denoted by S. Assume that $S = S(q, Q, t)$. From the earlier discussion, it follows that $p = \partial S/\partial q$:= Jacobi Law (Defn). Also, $K = H + \partial S/\partial t \implies \partial S/\partial t + H(q, \partial S/\partial q, t) = 0$:= Hamilton Jacobi Equation (Defn). This is a first order partial differential equation in $(n+1)$ variables, q, t. It is sufficient to define S. Defn: S :=Hamilton's principle function. Since only derivatives of S appear in the EOM, only n non-additive constants are relevant, viz. $S = S(q, \alpha, t)$. By assumption, $\alpha = Q$; however in general, one may use a combination of Qs and Ps.

Scheme 1: Consider the following overview of how this is implemented in practice. (1) Substitute $\partial S/\partial q = p$ in $\partial S/\partial t + H = 0$. (2) Solve for S and obtain n non-additive integration constants, $\alpha_1, \ldots \alpha_n$. (3) Define independent linear combinations of α_i as $Q_i \equiv \gamma_i(\alpha_1, \ldots \alpha_n)$. (4) Now recall that $P_i = -\partial S/\partial \gamma_i \equiv \beta_i$. From this equation, in principle, q can be obtained in terms of (β, γ), granted $\det\left(\partial^2 S/\partial q_i \partial \gamma_j\right) \neq 0$. [] Where did the Q, β get defined? Well, you specify it by matching the initial conditions t_0, q_0, p_0 with Q_i and $P_i (= \beta_i)$ using appropriate linear combinations of α_i. []

B.1.2 The Action Function

Assume that the system starts from (q_0, t_0) and reaches (q, t) through the path γ. Accordingly, the (extremized) action will be $I(q, t; q_0, t_0) = \int_\gamma L(q, \dot{q}, t) dt$. The idea is to look at variations in I, induced by varying q, t and q_0, while holding t_0 fixed. The motivation for this is the Hamilton Jacobi formalism, as was just set up. Upon substituting for H in the expression for I, one can evaluate $dI = pdq - Hdt - p_0 dq_0$. [] This step needs some work to derive, but take it as a claim for the moment [] Also, $dI = (\partial I/\partial q)dq + (\partial I/\partial t)dt + (\partial I/\partial q_0)dq_0$. Equating these yields $\partial I/\partial q = p$, $H = -\partial I/\partial t$, $p_0 = -\partial I/\partial q_0$. Using the first two, it surprisingly follows that I is a valid Hamilton's principle function. [] Ofcourse, now it's clear from the aforesaid equations that I generates canonical transformations which trivializes the motion ($K = 0$, nothing evolves). Here $q_0 \equiv Q$ and $p_0 \equiv P$ []

In Scheme 1 (Sec. B.1.1), the Hamilton Jacobi equation is integrated to solve Hamil-

ton's equation. The converse, viz. solving the Hamilton's equation to solve the Hamilton Jacobi equation, subject to $S(q,t_0) = S_0(q)$, can be achieved (it's a "Cauchy Problem"[1]); this is Scheme 2. $S(q,t_0) = S_0(q)$. (after this is not quite clear. Ask Arvind sir)

B.1.3 The Basic Law of Motion

There's an issue: if a complete solution of the Hamilton-Jacobi equation is not available, then the Jacobi Law can't be used. We assume that $S(q,t)$ is known, but the constants are not, viz. $\partial S/\partial \gamma_i = \beta_i$ doesn't help. Recall that $p_i = \partial S/\partial q_i = \partial L/\partial \dot{q}_i$. From L, one can evaluate $p_i = p_i(q_i,\dot{q}_i,t)$ and then use $p_i = \partial S/\partial \gamma_i$ to integrate for $q(t)$ with $q_0 = q(0)$.

Remarks: (1) For consistency, we must impose $p_0 = (\partial S/\partial q)_{q=q_0,t=t_0}$. (2) Even when the complete integral is known, p_i as a function of L is used to specify the initial conditions. (3) Since $p_i = \partial S/\partial q_i$ always holds, it maybe regarded as the basic law of motion.

Illustration (1) [using Scheme 1]: Recall that $p = \nabla S$. Assume $L = m\dot{q}^2/2 - V(q,t)$. $\implies$ $p = m\dot{q}$. S must satisfy $\partial S/\partial t + \nabla S^2/2m + V = 0$. For a free particle, we may use $S = P_i P_i t/2m + P_i q_i$ as a solution, where P_i is some constant. $P = \nabla S = p = m\dot{q}$. This integrates to $q = q_0 + P(t-t_0)/m$. We can also use the action, $S = m(q-q_0)^2/2t$ as a solution. [) That can be worked out as follows: Recall that $I = \int_\gamma L dt$. For a free particle, $= \int_\gamma (m\dot{q}^2/2)dt = \int_\gamma (1/2)m(dq/dt)(dq/dt)dt$. Now to impose γ as the path, we must put $dq/dt = v$ (constant). Thus, $I = mv(q-q_0)/2 = m(q-q_0)^2/2t$. It might appear pointless to solve the motion, when we have already evaluated γ. The point is that, given S, how to solve the motion. This was just one way of finding S. (] From $p = \nabla S$, we get $p = m(q-q_0)/t$; $P = \partial S/\partial Q$, with $Q = q_0$, yields $P = p$. Thus we recover $q = q_0 + P(t-t_0)/m$. Remarks: (1) S in the former case, is constant in q,t while in the latter, it explicitly depends on q,t. (2) S may not have a constant value along the particle trajectory; eg. $S = m(q-q_0)^2/2t = mv^2t^2/t$ for a free particle. (3) While S may not be single valued, physical quantities will be well defined, such as momentum (∇S), energy ($-\partial S/\partial t$) etc. [) Note how we started by saying that $S = S(q,Q,t)$ but suddenly, there's P in the first definition of S but not Q, though in the second definition, $S = S(q,Q,t)$ where $Q = q_0$. So how is the former a valid S (in the sense that how can we use results we derived by assuming $S(q,Q,t)$ as opposed to $S(q,P)$ (]

[Ask Arvind sir: super confusing] Let us now explore how..

Illustration (2): Consider $S(q,E,t) = W(q,E) - Et$

[1] A PDE problem with constraints defined on some hyper-surface of the domain

B.2 Densities

1. To understand the continuity equation $\rho + \nabla.(\rho v) = 0$, examples of v are taken

   (a) $v = v(x)$, a solution is obtained as

   $$\rho(x,t) = \frac{1}{v(x)} v \left[x \left(t - \int \frac{dx}{v} \right) \right] \rho_0 \left[x \left(t - \int \frac{dx}{v} \right) \right]$$

   in which further assuming that $\rho(x)$ results in $\rho = A/|v(x)|$

   i. $v = v(t)$ then we get

   $$\rho(x,t) = \rho_0 \left(x - \int v dt \right)$$

   which means that ρ is constant along particle trajectories

   (b) connection with Liuoville's equation

   i. $f(x,p,t)$ is defined instead of $\rho(x,t)$. Pure and mixed states are defined accordingly as $f(x,p,t) = \rho(x,t)\delta(p - \nabla S(x,t))$ being pure and the remaining as mixed.

   ii. $\frac{df}{dt} = \partial_t f + \frac{1}{m}\sum p_i \partial_{x_i} f - \sum \partial_{x_i} V \partial_{p_i} f = 0$ is the Liuoville's equation (which holds since we can show that the volume doesn't change under Hamiltonian evolution and particles inside the volume stay inside; $f(p',q',t+\delta t) = f(p,q,t)$ is essentially the statement $\frac{df}{dt} = 0$) which is linear in f.

   iii. One may project out the moment space. They define equivalent of ρ as $P(x) = \int f d^3p$, mean momentum as $\overline{p_i(x)} = \frac{\int p_i f d^3p}{P(x)}$ and $\overline{p_i p_j(x)} = \frac{\int p_i p_j f d^3p}{P(x)}$. The louvielle equation can then be expressed in terms of these spatial variables. Integrating it we get

   $$\partial_t P + \frac{1}{m}\sum \partial_{x_i}(P\overline{p_i}) = 0.$$

   To get the momentum transport equation, after multiplying the louvielle equation with p_i and integrating, we get

   $$\partial_t(P\overline{p_i}) + \frac{1}{m}\sum \partial_{x_j}(P\overline{p_i p_j}) + P\partial_{x_i} V = 0$$

   (apparently integrated by parts and assumed $f \to 0$ as $p_i \to \infty$)
   While f is constant along a phase space trajectory, the spatial density P (equivalent of ρ) is not. It's apparent from the derivation of the continuity equation; either we start with a fixed volume or a fixed number of particles, not both.
   If you substitute $f = \rho\delta(p - \nabla S)$ as stated earlier, you'd get $P = \rho$, $\overline{p_i} = \partial_{x_i} S$,

$\overline{p_i p_j} = \partial_{x_i} S \partial_{x_i} S$ as expected. The substitution also yields what's called a field theoretic version of Newton's Laws given by

$$\partial_t \rho + \frac{1}{m} \nabla . (\rho \nabla S) = 0$$

and

$$\left[\partial_t + \frac{1}{m} \sum \partial_{x_i} S \partial_{x_j} \right] \partial_{x_i} S = 0$$

iv. Remarks:

A. It's not obvious that if we start with a state that has well defined momentum (delta distribution) but the positions are given by $\rho(x)$, then they will continue to be well defined in momentum. This happens only exceptionally. In general, a pure state maybe sent to a mixed state. We'll see examples of these. [todo: ensure examples make sense]

B. Can we decompose any mixed ensamble into a linear combination of pure ones? The answer's no. [proof?] Say there are many solutions of the Hamilton-Jacobi equation, given by S_i. Thus, we can construct a linear combination as $f(x,p,t) = \sum P_i \rho_i(x,t) \delta(p - \nabla S_i(x,t))$ where P_i (degenerate notation) refers to the distribution of momenta at a given point. $\sum P_i = 1$ is assumed for normalization. Claim is that this is not in general possible to decompose a state into this form. An explicit example is that of reflecting through a potential barrier (in CM) [todo: ensure the example works]

C. While this is not particularly useful in CM (the pure and mixed states), the formalism helps in comparison with QM.

(c) Pure and Mixed States

i. Illustration: We see that $f_0(x,p) = \delta(x - x_0)\delta(p - p_0)$ remains sharp
(it can be checked by inserting it in the louviel equation) to yield $f(x,p) = \delta(x - x(t,x_0,p_0))\delta(p - p(t,x_0,p_0))$ [this is expected, since you're in essence saying there's only one particle]

ii. Illustration 2: We want to see what happens to a Gaussian like state, does it spread?
We start with $\rho_0 = \frac{e^{-x^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$ and $S_0 = px$ with σ and p constant. This form of S_0 has already been solved for and tells us $\rho = \frac{e^{-(x-vt)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$. There's no spreading classically! We'll see for the same initial conditions, what happens quantum mechanically.

iii. Illustration 3: What initial conditions yield a spreading Guassian? We start with the same ρ_0 but use $S_0 = \frac{m(x-x_0)^2}{2t}$, in which case the solution we saw the result is

B.3 Aside

So we start with $\oint dS = \oint \nabla S.dx = \oint p.dx = \int \nabla \times p.da$. If $\oint dS = nh$, then we must have $\nabla \times p = \sum_a \Gamma_a \int_{\gamma_a} \delta(x - x_a) dx_a$ where γ_a is the nodal line. If we assume $\oint dS = nh$ holds, then can we construct some example of the same? Let's first see how $\oint dS = nh$ can be derived. If the only condition is that ψ is single valued, then we know that at any point, S' and S both yield the same ψ, where $S' = S + nh$. If one considers a loop, then say we start from a point S_a. Then after completing some distance, the change in S is given by ΔS. So the value of S starting from $S_a$will be $S_a + \Delta S$. Now if we come back to the point a, then from uniqueness of ψ, we only demand $S_a + \Delta S = S_a + nh$. If S itself was unique, then we'd say $S_a + \Delta S = S_a$. Now at this point itself I seem to have trouble. I have tacitly assumed that S is single valued when I'm evaluating the 'change is S' along the curve.

Talked to manu for a while and made some progress, then figured it was non-sense and made some more progress. Finally, Manu found a document that helped clarify a few things. The issue was still that they had used a vector field and not a potential. And it wasn't clear to me what potential must I use in that case.

The potential is $V = k\theta$. Note how this is itself, as a function of position is multi valued and yet we never have any issues integrating this (as we'll see shortly). While V is multivalued, $\nabla V = \frac{k}{r}\hat{\theta}$ is happily single valued :) And not just that, check this; $\oint_\gamma \nabla V.dx = 2\pi$ (simply because γ is chosen to be a circle and then $dx = rd\theta\hat{\theta}$). Since in the domain of interest, everything is well defined, I can write $\oint_\gamma \nabla V.dx = \oint_\gamma dV = 2\pi$. And one can show independently (I know only a simple minded proof with discritizing the function) that $\oint dV = 0$ whenever V is single valued (or a function). So what does this example show? Various rather peculiar things. (I) that $\oint dV$ maybe non zero for a reasonable physical situation by virtue of multivaluedness of V. Yes, V is multivalued and yet we can integrate the said expression without ambiguity. (II) that there happens to be a singularity within the loop, over which the integral is non-zero. (III) The curl, $\nabla \times \nabla V \neq 0$ at the center and $= 0$ else.
Now we've made plausible various things which would've seemed arbitrary otherwise.