

An alternative to contextuality

Atul Singh Arora^{1,*} and Arvind^{1,†}

¹*Indian Institute of Science Education and Research (IISER), Mohali*

We propose a non-contextual hidden variable model, consistent with all predictions of QM (for discrete spaces). We reconsider the no-go theorems on hidden variables and conclude that the notion of contextuality is not a necessary feature of quantum mechanics. Motivated by consistency requirements between Bohmian Mechanics and the said theorems, we provide an alternative view. This is accomplished by identifying a new classical property, we call ‘multiplicativity’. Advantages of this view are illustrated by considering the implications of non-multiplicativity to non-locality and its relation to contextuality.

* msl1003@iisermohali.ac.in

† arvind@iisermohali.ac.in

I. INTRODUCTION

Background: Since Einstein's [1] work on completeness of quantum mechanics and the subsequent seminal work of Bell, assessing compatibility of hidden variable models has proven to be repeatedly useful. It has had foundational implications, for instance it gave a precise meaning to locality, and engendered pragmatic advantages, such as security, randomness certification etc. While Bell showed that local hidden variable models can not describe nature, the Kochen Specker theorem [2] (and related work [3–6]) ruled out an even larger class of hidden variable models. The notion of contextuality was identified which also, in recent times has been harnessed for computation and cryptography. Not all hidden variable models were, however, proved to be incompatible. Bohm's work [7, 8] is one such (in)famous example. This approach has allowed for a precise interpretation of quantum mechanics.

Previous work [the essentials I have already discussed]: In this paper, we take a closer look at the aforesaid no-go theorems and find that contextuality is in fact not a necessary feature of quantum mechanics. This has been shown [9] by taking disturbances due to measurements into consideration. Here, however, we construct a significantly simpler toy model for illustrating not only the aforesaid but to also delineate a complete alternative to contextuality. The toy model is a specific case of what we call a c-angle variable theory. Its predictions match with those of quantum mechanics for systems described by arbitrary (discrete) Hilbert spaces.

Triangle: <TODO: ask arvind sir>

Work and results: <summarise what is done in each section>

II. CONTEXT

The notation used henceforth has been defined below.

Definition 1. Notation: (a) $\psi \in \mathcal{H}$ would typically represent the quantum mechanical state of the system (assumed pure), residing in the Hilbert space \mathcal{H} , (b) $\hat{\mathcal{H}}$ is defined to mean $\mathcal{H} \otimes \mathcal{H}^\dagger$, (c) $[\mathcal{H}]$ is defined to mean $(\mathcal{H}, \mathbb{R}^\otimes)$, which represents the state of the system including hidden variables, (d) $[\psi] \in [\mathcal{H}]$ will represent the state of the system, including hidden variables, (e) a prediction map is $m : \hat{\mathcal{H}}, [\mathcal{H}] \rightarrow \mathbb{R}$, (f) a sequence map is $s : \hat{\mathcal{H}}, [\mathcal{H}], \mathbb{R} \rightarrow [\mathcal{H}]$, (g) the set of all experimental setups is denoted by \mathcal{E} , (h) a setup map is $e : \hat{\mathcal{H}} \rightarrow \mathcal{E}$.

In addition to the notation, we use the following definition of non-contextuality.

Definition 2. A theory is non-contextual, if it provides a map $m : \hat{\mathcal{H}}, [\mathcal{H}] \rightarrow \mathbb{R}$ to explain measurement outcomes. A theory which is not non-contextual is contextual.

Remark 3. Prediction maps are non-contextual.

In the literature broader definitions have been suggested which declare a larger set of theories as non-contextual. For our purposes however, this restricted definition will suffice. The term contextual is used to suggest that the value an operator takes might depend on which other compatible observable it is measured with. Since quantum mechanics (QM) yields deterministic results only in specific situations, it is imperative to analyse the meaning of contextuality by ‘completing’ QM. Hidden variable theories do precisely this and we have at our disposal the heretic theory of Bohm. Circumventing the details of the theory, we note its salient feature; with each particle, one associates precisely defined (q, p) and a wavefunction ψ . Consequently, the outcomes of all experiment are predictable, granted the initial conditions of the system and the apparatus are known (the ‘hidden variables’). The essence of Bohmian Mechanics relevant here can be captured by the following theorem.

Theorem 4. \exists a (Bohmian) map $m_B : \hat{\mathcal{H}}, [\mathcal{H}], \mathcal{E} \rightarrow \mathbb{R}$, and a sequence map s , s.t. if m_B & s are assumed to describe the outcomes of measurements & the resultant state respectively, then they are consistent with all predictions of QM.

Observe that given a Stern-Gerlach setup and the initial conditions (position & wavefunction of the particle), according to the theorem, Bohmian mechanics can predict the value of the spin (say $\hat{\sigma}_z$) of the particle. Assume that the state of the particle was $(|0\rangle + |1\rangle)/\sqrt{2}$ and that the measurement outcome was spin up. Interestingly, it can be shown [10] that if the direction of the magnetic field gradient is flipped while everything else is unchanged, the measurement outcome will now be spin down. This emphasises the fact that m_B is not a prediction map. More importantly it entails that m_B is contextual. It must be stressed, however, that the contextuality in this case is not a statement about the context in terms of compatible observables. Here no two observables are compatible (excluding identity).

III. NON-MULTIPLICATIVITY: AN ALTERNATIVE TO CONTEXTUALITY

One can show [2] that under reasonable assumptions, no non-contextual theory can explain all predictions of QM. We re-state these results and show that nowhere does contextuality appear to be a logical necessity. Instead we note that non-contextual theories with certain properties, we call *multiplicativity* are incompatible with QM. Viewed from this perspective, non-*multiplicativity* is the non-classical alternative to contextuality in QM. Precise definitions of this and a related property *sequential multiplicativity*, are given below.

Definition 5. A prediction map m is *multiplicative* iff

$$m(f(\hat{B}_1, \hat{B}_2, \dots \hat{B}_N), [\psi]) = f(m(\hat{B}_1, [\psi]), m(\hat{B}_2, [\psi]), \dots m(\hat{B}_N, [\psi])),$$

where $\hat{B}_i \in \hat{\mathcal{H}}$ are arbitrary mutually commuting observables, $f : \hat{\mathcal{H}}^{\otimes N} \rightarrow \hat{\mathcal{H}}$ and $[\psi] \in [\mathcal{H}]$. A non-*multiplicative* map is one that is not multiplicative.

Note that if m is taken to represent the measurement outcome (in QM), then for states of the system which are simultaneous eigenkets of \hat{B}_i s, m must clearly be multiplicative. It is, however, not obvious that this property must always hold. To motivate this, consider the state $|11\rangle := |1\rangle \otimes |1\rangle$ written in the computational basis, $\hat{B}_1 = \hat{\sigma}_x \otimes \hat{\sigma}_x$, $\hat{B}_2 = \hat{\sigma}_y \otimes \hat{\sigma}_y$ and $\hat{C} = \hat{B}_1 \hat{B}_2 = -\hat{\sigma}_z \otimes \hat{\sigma}_z$, we have $m(\hat{C}) = -1$ while $m(\hat{B}_1) = \pm 1$ and $m(\hat{B}_2) = \pm 1$, independently, according to QM, with probability half. Contrary to this, from a quick computation, it is clear that if one first measures \hat{B}_1 and subsequently measures \hat{B}_2 , then the product of the results must be -1 . This is consistent with measuring \hat{C} . To capture this property we define *sequential multiplicativity* as follows.

Definition 6. A prediction map m is *sequentially multiplicative* for a given sequence map s , iff

$$m(f(\hat{B}_1, \hat{B}_2, \dots \hat{B}_N), [\psi_1]) = f(m(\hat{B}_1, [\psi_{k_1}]), m(\hat{B}_2, [\psi_{k_2}]), \dots, m(\hat{B}_N, [\psi_{k_N}])),$$

where $\mathbf{k} = (k_1, k_2, \dots, k_N) \in \{(1, 2, 3 \dots N), (2, 1, 3 \dots N) + \text{permutations, making } N! \text{ terms}\}$, $\hat{B}_i \in \hat{\mathcal{H}}$ are arbitrary mutually commuting observables, $[\psi_i] \in [\mathcal{H}]$, $f : \hat{\mathcal{H}}^{\otimes N} \rightarrow \hat{\mathcal{H}}$ and $[\psi_{k+1}] \equiv s(\hat{B}_k, [\psi_k], m(\hat{B}_k, [\psi_k]))$, $\forall [\psi_1]$.

We now discuss the ‘proofs of contextuality’ in the purview of the aforesaid.

Theorem 7. Let a map $m : \hat{\mathcal{H}} \rightarrow \mathbb{R}$, be s.t. (a) $m(\hat{\mathbb{I}}) = 1$, (b) $m(f(\hat{A}_1, \hat{A}_2, \dots)) = f(m(\hat{A}_1), m(\hat{A}_2), \dots)$, for any arbitrary function f , where \hat{A}_i are arbitrary Hermitian operators. If m is assumed to describe the outcomes of measurements, then no m exists which is consistent with all predictions of Quantum Mechanics.

Proof. Bell construction ($|\mathcal{H}| \geq 4$): Consider an operator $\hat{B} := a_1 \otimes b_1 + a_2 \otimes b_1 + a_1 \otimes b_2 - a_2 \otimes b_2$ s.t. $-1 \leq \hat{a}_i, \hat{b}_j \leq 1 \forall i, j \in \{1, 2, 3, 4\}$ (which means that the spectrum of \hat{a} and \hat{b} are bounded above and below). If we assume that a map m of the aforesaid type exists, then it can be shown that $\langle \hat{B} \rangle \leq 2$. To see this, consider $\hat{A}_1 = \hat{a}_1 \otimes \hat{\mathbb{I}}$ and $\hat{A}_2 = \hat{\mathbb{I}} \otimes \hat{b}_1$, so that if $\hat{C} := \hat{A}_1 \hat{A}_2$ then property (2) entails that $m(\hat{C}) = m(\hat{A}_1)m(\hat{A}_2)$, viz. $m(\hat{a}_1 \otimes \hat{b}_1) = m(\hat{a}_1 \otimes \hat{\mathbb{I}})m(\hat{\mathbb{I}} \otimes \hat{b}_1)$. Note that $m(\hat{*})$, may depend on the state $|\psi\rangle$ and the hidden variables. Therefore, we average over multiple measurements, which corresponds to averaging over the hidden variables. This is expressed by $\langle m(\hat{a}_1 \otimes \hat{b}_1) \rangle = \langle m(\hat{a}_1 \otimes \hat{\mathbb{I}}) \rangle \langle m(\hat{\mathbb{I}} \otimes \hat{b}_1) \rangle$. One can repeat this argument for each term in $\langle \hat{B} \rangle$. Using the fact that $-1 \leq m(\hat{*}) \leq 1$, and that the extrema will occur only at the extreme values of m , one can plugin ± 1 for each $\langle m(\hat{*}) \rangle$ to obtain $\langle \hat{B} \rangle \leq 2$. Quantum mechanically, one can show that $\langle \hat{B} \rangle = 2\sqrt{2}$ for 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, given in the z-basis as

$$\hat{\sigma}_x \doteq \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \hat{\sigma}_y \doteq \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \hat{\sigma}_z \doteq \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The observables are $\hat{a}_1 = \hat{\sigma}_z$ and $\hat{a}_2 = \hat{\sigma}_x$, while $\hat{b}_1 = -\frac{\hat{\sigma}_z + \hat{\sigma}_x}{\sqrt{2}}$ and $\hat{b}_2 = \frac{\hat{\sigma}_z - \hat{\sigma}_x}{\sqrt{2}}$. Since quantum mechanics violates the inequality, $\langle \hat{B} \rangle \leq 2$, we conclude that the map m can not simultaneously exist and be consistent with QM. \square

Proof. GHZ construction ($|\mathcal{H}| \geq 6$): Consider three observers with one particle each. Each observer can measure two properties, call them X and Y, with outcomes ± 1 . The associated operators are $\hat{\sigma}_x$ and $\hat{\sigma}_y$. The state of the system is $\sqrt{2} |\chi_G\rangle = |000\rangle - |111\rangle$ and we also define $\hat{A} := \hat{\sigma}_x \otimes \hat{\sigma}_y \otimes \hat{\sigma}_y$, $\hat{A} |\chi_G\rangle = |\chi_G\rangle$, $\hat{B} := \hat{\sigma}_y \otimes \hat{\sigma}_x \otimes \hat{\sigma}_y$ and $\hat{C} := \hat{\sigma}_y \otimes \hat{\sigma}_y \otimes \hat{\sigma}_z$. A measurement of \hat{A} would correspond to measurement of property X, Y and Y by the three observers respectively. For a map m consistent with measurement outcomes of QM, we must have

$m(\hat{A}) = m(\hat{B}) = m(\hat{C}) = +1$. Property (b) of the map $\implies 1 = m(\hat{A}\hat{B}\hat{C}) = m(\hat{\sigma}_x \otimes \hat{\sigma}_y \hat{\sigma}_x \hat{\sigma}_y \otimes \hat{\sigma}_x) = m(\hat{\sigma}_x^{(1)})m(\hat{\sigma}_y^{(2)}\hat{\sigma}_x^{(2)}\hat{\sigma}_y^{(2)})m(\hat{\sigma}_x^{(3)}) = m(\hat{\sigma}_x^{(1)})m(\hat{\sigma}_x^{(2)})m(\hat{\sigma}_x^{(3)}) = m(\hat{D} \equiv \hat{\sigma}_x \otimes \hat{\sigma}_x \otimes \hat{\sigma}_x)$, where $\hat{\sigma}_x^{(1)} \equiv \hat{\sigma}_x \otimes \hat{\mathbb{I}} \otimes \hat{\mathbb{I}}$ and so on. It must be stressed that although \hat{A} , \hat{B} and \hat{C} mutually commute, $\hat{\sigma}_x^{(2)}, \hat{\sigma}_y^{(2)}$ do not. We used property (b) over both these sets of operators in the previous step. Returning to the argument, note that $\hat{D}|\chi_G\rangle = -|\chi_G\rangle$, $\implies m(\hat{D}) = -1$ which yields a contradiction. We therefore conclude that no m with the said properties exists which is consistent with all predictions of QM. \square

Here m maybe viewed as a specific class of prediction maps, that implicitly depends on the state $[\psi]$. It follows then that the maps ruled out by the theorem are non-contextual. The set of maps ruled out, however, can be enlarged by imposing fewer restrictions on it. This is achieved by the following theorem.

Theorem 8. *Let a map $m : \hat{\mathcal{H}} \rightarrow \mathbb{R}$, be s.t. (a) $m(\hat{\mathbb{I}}) = 1$, (b) $m(f(\hat{B}_1, \hat{B}_2, \dots)) = f(m(\hat{B}_1), m(\hat{B}_2), \dots)$, for any arbitrary function f , where \hat{B}_i are mutually commuting Hermitian operators. If m is assumed to describe the outcomes of measurements, then no m exists which is consistent with all predictions of Quantum Mechanics.*

Proof. Bell construction ($|\mathcal{H}| \geq 4$): The same proof as the one presented earlier goes through here also since $[A_1, A_2] = 0$. \square

Proof. Generalized GHZ construction ($|\mathcal{H}| \geq 6$): Consider again, three observers with one quantum particle each. Each observer can measure the properties corresponding to $\hat{\sigma}_x$ and $\hat{\sigma}_y$ of his/her particle. The quantum state of these particles is also the same as before, $\sqrt{2}|\chi_G\rangle = |000\rangle - |111\rangle$. We define as earlier $\hat{A} := \hat{\sigma}_x \otimes \hat{\sigma}_y \otimes \hat{\sigma}_y \implies \hat{A}|\chi_G\rangle = |\chi_G\rangle$, $\hat{B} := \hat{\sigma}_y \otimes \hat{\sigma}_x \otimes \hat{\sigma}_y$ and $\hat{C} := \hat{\sigma}_y \otimes \hat{\sigma}_y \otimes \hat{\sigma}_x$. Consider in addition the following operators, represented by \hat{H}_{ij} .

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

where from property (b) it follows that $m(\hat{A}) = m(\hat{H}_{11})m(\hat{H}_{12})m(\hat{H}_{13})$. Note that $[\hat{H}_{ij}, \hat{H}_{ik}] = 0$, viz. the operators along a row commute. To be consistent with QM, $m(\hat{A}) = +1$ (similarly for row 2 and 3). $m(\hat{D}) = -1$ imposes that the product $m(\hat{H}_{41})m(\hat{H}_{42})m(\hat{H}_{43})$, row 4, must be -1 . Using the notation $m(\hat{H}_{ij}) = H_{ij}$, we can start with the first case. For the last row to yield -1 , we must have either one -1 assignments or three -1 assignments. Let us look at the first case. Without loss of generality, we can assume that $H_{41} = H_{42} = +1$ and $H_{43} = -1$. Further, $H_{11} = H_{41}$, $H_{22} = H_{42}$ and $H_{33} = H_{43}$ by construction (see below).

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

Next, one must ensure that the first row yields $+1$ which requires us to either set $H_{12} = H_{13}$ to -1 or $+1$. Let us write this as $H_{12} = H_{13} = \pm 1$ which would entail that $H_{23} = \pm 1$ and $H_{32} = \pm 1$. [Note that the notation implies that when we choose $+1$, all these are $+1$ and when we choose -1 , all are -1 . We are not allowed to set, say $H_{12} = -1$ and $H_{13} = +1$.] For the second row to be $+1$, we conclude that $H_{21} = \pm 1$. This entails, in turn, that $H_{31} = \pm 1$. We now note that the third row has become -1 whereas according to QM, it must be $+1$. It is left upto the reader to see that similar reasoning also yields results in inconsistencies with QM. This completes the proof. \square

The following proof is not only more general, but also independent of the initial state of the system.

Proof. Peres Mermin ($|\mathcal{H}| \geq 4$):

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 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 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 \hat{R}_i and \hat{C}_j , it is easy to show that no map m with the stated properties can exist. Let us assume on the contrary that m exists. From property (b) of the map, to get $m(\hat{C}_3) = -1$ (used property (a)), 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 $m(\hat{R}_i) = 1$, 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 a consistent m exists, must be wrong. \square

Proof. Kochen Specker ($|\mathcal{H}| \geq 3$): The original proof given by [2] is rather involved. Here we will satisfy ourselves with assuming its validity and applying our constructions to the aforesaid proof(s). \square

Remark 9. One could in principle assume m , to be s.t. (a) $m(\hat{\mathbb{I}}) = 1$, (b) $m(\alpha\hat{B}_i) = \alpha m(\hat{B}_i)$, for $\alpha \in \mathbb{R}$, (c) $m(\hat{B}_i^2) = m(\hat{B}_i)^2$, (d) $m(\hat{B}_i + \hat{B}_j) = m(\hat{B}_i) + m(\hat{B}_j)$, to deduce (d') $m(\hat{B}_i\hat{B}_j) = m(\hat{B}_i)m(\hat{B}_j)$ and that $m(\hat{B}_i) \in$ spectrum of \hat{B}_i . Effectively then, condition (b) listed in the theorem is satisfied as a consequence. Therefore, assuming (a)-(d) as listed here, rules out a larger class of m .

According to theorem 8, non-contextual maps which are *multiplicative* must be incompatible with QM. It is apparent that non-contextual maps which are non-*multiplicative* might be consistent with QM. To that end, in the following section we first show that this is indeed the case by removing contextuality from Bohmian Mechanics and supplement it thereafter by an explicit construction.

Before proceeding we note, however, that QM enforces *sequential multiplicativity* at least in certain cases. This result will be used to assess accuracy of any hidden variable model we might propose.

Proposition. *Let the system be in a state, s.t. measurement of \hat{C} yields repeatable results (same result each time). Then according to QM, sequential multiplicativity holds, where $\hat{C} \equiv f(\hat{B}_1, \hat{B}_2, \dots, \hat{B}_n)$, and \hat{B}_i are as defined.*

Proof. Assume without loss of generality that $\hat{B}_1, \hat{B}_2, \dots, \hat{B}_n$ are mutually compatible (commuting) and complete set of operators. If say for instance the set is not complete, then one can add the missing operators and label them as aforesaid. It follows that $\exists |\mathbf{b}\rangle = \left| b_1^{(l_1)}, b_2^{(l_2)}, \dots, b_n^{(l_n)} \right\rangle$ s.t. $\hat{B}_i |\mathbf{b}\rangle = b_i^{(l_i)} |\mathbf{b}\rangle$, where l_i indexes the eigenvalues corresponding to \hat{B}_i and that $\sum_{\mathbf{b}} |\mathbf{b}\rangle \langle \mathbf{b}| = \hat{\mathbb{I}}$. Let the state of the system be given by $|\psi\rangle$ and it must be s.t. $\hat{C} |\psi\rangle = c |\psi\rangle$, by assumption. For the statement to follow, one need only show that $|\psi\rangle$ must be made of only those $|\mathbf{b}\rangle$ s, which satisfy $c = f(b_1^{(l_1)}, b_2^{(l_2)}, \dots, b_n^{(l_n)})$. This is the crucial step. Proving this is so is trivial. We start with $\hat{C} |\psi\rangle = c |\psi\rangle$ and take its inner product with $\langle \mathbf{b}|$ to get

$$\begin{aligned} \langle \mathbf{b}| \hat{C} |\psi\rangle &= c \langle \mathbf{b}| \psi\rangle, \\ \langle \mathbf{b}| f(\hat{B}_1, \hat{B}_2, \dots, \hat{B}_n) |\psi\rangle &= c \langle \mathbf{b}| \psi\rangle, \\ f(b_1^{(l_1)}, b_2^{(l_2)}, \dots, b_n^{(l_n)}) \langle \mathbf{b}| \psi\rangle &= c \langle \mathbf{b}| \psi\rangle. \end{aligned}$$

Also, we have $|\psi\rangle = \sum_{\mathbf{b}} \langle \mathbf{b}| \psi\rangle |\mathbf{b}\rangle$, from completeness. If we consider $|\mathbf{b}\rangle$ s for which $\langle \mathbf{b}| \psi\rangle \neq 0$, then we can conclude that indeed $c = f(b_1^{(l_1)}, b_2^{(l_2)}, \dots, b_n^{(l_n)})$. However, when $\langle \mathbf{b}| \psi\rangle = 0$, viz. $|\mathbf{b}\rangle$ s that are orthogonal to $|\psi\rangle$, then nothing can be said. We can thus conclude that $|\psi\rangle$ is made only of those $|\mathbf{b}\rangle$ s that satisfy the required relation. That completes the proof. \square

It is worth noting that in the PM case, where \hat{R}_i and \hat{C}_j are just $\pm\hat{\mathbb{I}}$, it follows that all states are their eigenstates. Consequently, for these operators *sequential multiplicativity* must always hold.

A. Non-contextual Bohmian Mechanics

Recall that contextuality in Bohmian Mechanics arises from the mapping between an operator and the experimental setup. Consider the following proposition.

Lemma 10. *For discrete \mathcal{H} , \exists a map $e : \hat{\mathcal{H}} \rightarrow \mathcal{E}$.*

Proof. To describe an experimental setup corresponding to a measurement, we use a measuring particle (mass m_L , say). It is made to interact with the system for a short duration appropriately and subsequently we measure the resultant 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 $\hat{H}_{\text{int}} = a\hat{L} \otimes \hat{p}$. Here the first operator acts on the system and second (after the tensor product symbol) acts on the measuring particle. a quantifies the interaction strength. [TODO: check this] It has dimensions of frequency, if L has dimensions of length. Let us assume that the system is 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-degeneracy for simplicity, but its removal doesn't cause any significant difficulty). $|\Psi_S(t)\rangle = \hat{U}(t) |\psi\rangle \otimes |\varphi\rangle$, where

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

and $|\varphi\rangle$ is the state of the measuring particle, given by a Gaussian 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 $\hat{H}_{\text{int}} t$. We then have

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

where $|\varphi_{q_0}\rangle = \int dq \varphi(q - q_0) |q\rangle$. 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. Therefore, given an observable, one can construct the aforesaid experiment to measure its value. We have, in effect, explicitly constructed the map $e : \mathcal{H} \rightarrow \mathcal{E}$. \square

The proof of this proposition can also be motivated by a simple construction. Consider that the system of interest are spin half particles. To measure spin along any direction, there are two directions along which the magnetic field gradient can be set (as was discussed in the introduction). The experimentalist agrees to always use a specific direction, thereby mapping an operator to an experimental setup uniquely. Given this, one can immediately derive a consistent non-multiplicative prediction map m .

Proposition 11. *The Bohmian map, m_B can be restricted to a prediction map m , we call ‘Bohmian prediction map’.*

Proof. $m(\hat{\mathcal{H}}, [\mathcal{H}]) = m_B(\hat{\mathcal{H}}, [\mathcal{H}], e(\hat{\mathcal{H}}))$. \square

Proposition 12. *The Bohmian prediction map, m must be non-multiplicative.*

Proof. Since Bohmian Mechanics is consistent with all predictions of QM, the Bohmian prediction map m must also be. According to theorem 8, m must be non-multiplicative. \square

B. C-angle model | An Explicit Construction

For readers skeptic about Bohmian Mechanics, the discussion related to it can be replaced by the following. Let the state of the system be $|\chi\rangle$, defined on a discrete Hilbert space (spin like) and we wish to assign a value to an arbitrary operator $\hat{A} = \sum_a a |a\rangle \langle a|$, which has eigenvalues $\{a_{\min} = a_1 \leq a_2 \leq \dots \leq a_n = a_{\max}\}$. This model has the following postulates:

1. Initial Hidden Variable: Pick a $c \in [0, 1]$, from a uniform random distribution.
2. Assignment/Prediction: The value assigned to \hat{A} is given by finding the smallest a s.t.

$$c \leq \sum_{a'=a_{\min}}^a |\langle a'|\chi\rangle|^2,$$

viz. we have specified a prediction map, $m(\hat{A}) = a$.

3. Update: After measuring an operator, the state must be updated (collapsed) in accordance with the rules of QM. This completely specifies the sequence map s .

To see how this works, we restrict ourselves to a single spin case. Say $|\chi\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$, and $\hat{A} = \hat{\sigma}_z = |0\rangle\langle 0| - |1\rangle\langle 1|$. Now, according to the postulates of this theory, $m(\hat{A}) = +1$, if $c \leq \cos^2\theta$, else \hat{A} will be assigned -1 . It follows then, from c being uniformly random in $[0, 1]$, that the statistics agree with predictions of QM. The reader can convince him(her)self that the said scheme works in general. A more interesting application of this to the PM situation follows.

We can clearly see that the assignment, described by the prediction map m , is non-contextual since given an operator and a state (+ the hidden variable), the value is uniquely assigned. The map m is non-multiplicative which will be clarified by the following.

Example 13. [C-angle model applied to the Peres Mermin situation] We apply the C-angle model to the Peres Mermin situation discussed earlier. For illustration, let us start with the state $|\psi_1\rangle = |00\rangle$. Say from a uniform random distribution we obtained $c = 0.4$. To arrive at the assignments, note that $|00\rangle$ is an eigenket of only \hat{R}_i, \hat{C}_j and $\hat{A}_{33} = \hat{\sigma}_z \otimes \hat{\sigma}_z$. Thus, in the first iteration, all these should be assigned their respective eigenvalues. The remaining operators must be assigned -1 as one can readily verify by explicitly finding the smallest a as described in postulate 2 of the model (see equation (1)). Two remarks are in order. First, this model is manifestly *non-multiplicative*, for $m_1(\hat{C}_3) = 1 \neq m_1(\hat{A}_{13})m_1(\hat{A}_{23})m_1(\hat{A}_{33}) = -1$, where the subscript has been introduced to index the iteration. More precisely, $m_1(\hat{O}) := m(\hat{O}, [|\psi_1\rangle = |00\rangle])$ where the complete state $[|\psi_1\rangle]$ implicitly refers to both the quantum state $|00\rangle$ and the hidden variable $c = 0.4$. Second, we must impose *sequential multiplicativity* as a consistency check of the model, which in particular entails that $m_1(\hat{C}_3) = m_1(\hat{A}_{33})m_2(\hat{A}_{23})m_3(\hat{A}_{13})$, where $m_2 := m(\hat{O}, [|\psi_2\rangle])$, $m_3 := m(\hat{O}, [|\psi_3\rangle])$ and $|\psi_2\rangle, |\psi_3\rangle$ are obtained from postulate 3. Note that for each iteration, a new hidden variable is generated. To illustrate *sequential multiplicativity*, we must choose to measure \hat{A}_{33} . According to step 4, since $|00\rangle$ is an eigenstate of \hat{A}_{33} , the final state remains $|00\rangle$.

Iteration	$i = 1$	$i = 2$	$i = 3$
$ \psi_{\text{init}}\rangle$	$ 00\rangle$	$ 00\rangle$	$\frac{ 00\rangle + 11\rangle}{\sqrt{2}}$
HV/Toss	$c = 0.4$	$c = 0.1$	$c = 0.7$
Predictions	$m_1(\hat{A}_{ij}) \doteq \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & +1 \end{bmatrix}$	$m_2(\hat{A}_{ij}) \doteq \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & +1 \end{bmatrix}$	$m_3(\hat{A}_{ij}) \doteq \begin{bmatrix} +1 & +1 & +1 \\ +1 & +1 & -1 \\ +1 & +1 & +1 \end{bmatrix}$
(Assignments)	$m_1(\hat{R}_i), m_1(\hat{C}_j) = +1 (j \neq 3)$ $m_1(\hat{C}_3) = -1$	$m_2(\hat{R}_i), m_2(\hat{C}_j) = +1 (j \neq 3)$ $m_2(\hat{C}_3) = -1$	$m_3(\hat{R}_i), m_3(\hat{C}_j) = +1 (j \neq 3)$ $m_3(\hat{C}_3) = -1$
Operator			
Measured	$\hat{A}_{13} = \hat{\sigma}_z \otimes \hat{\sigma}_z; m_1(\hat{A}_{13}) = +1$	$\hat{A}_{23} = \hat{\sigma}_y \otimes \hat{\sigma}_y; m_2(\hat{A}_{23}) = -1$	$\hat{A}_{33} = \hat{\sigma}_x \otimes \hat{\sigma}_x; m_3(\hat{A}_{33}) = +1$
$ \psi_{\text{final}}\rangle$	$ 00\rangle$	$\frac{ 00\rangle + 11\rangle}{\sqrt{2}}$	$\frac{ 00\rangle + 11\rangle}{\sqrt{2}}$

(1)

For the next iteration, $i = 2$, viz. after the first measurement has been performed, say the random number generator yielded $c = 0.1$. Since $|\psi\rangle$ is also unchanged the assignment remains invariant (in fact any of $c < 0.5$ would yield the same result), which should be evident from the previous exercise. For the final step we choose to measure $\hat{p} = \hat{A}_{23} (= \hat{\sigma}_y \otimes \hat{\sigma}_y)$, to proceed with sequentially measuring \hat{C}_3 . To simplify calculations, we note

$$|00\rangle = \frac{(|\tilde{+}\tilde{-}\rangle + |\tilde{-}\tilde{+}\rangle)/\sqrt{2} + (|\tilde{+}\tilde{+}\rangle + |\tilde{-}\tilde{-}\rangle)/\sqrt{2}}{\sqrt{2}},$$

where $|\tilde{\pm}\rangle = |0\rangle \pm i|1\rangle$ (eigenkets of $\hat{\sigma}_y$). Since $|00\rangle$ is manifestly not an eigenket of \hat{p} , we must find an appropriate eigenket $|p_{-}\rangle$ s.t. $\hat{p}|p_{-}\rangle = -|p_{-}\rangle$, since $c = 0.1$ and $\langle p_{-}|00\rangle$ is already > 0.1 . It is immediate that $|p_{-}\rangle = (|\tilde{+}\tilde{-}\rangle + |\tilde{-}\tilde{+}\rangle)/\sqrt{2} = (|00\rangle + |11\rangle)/\sqrt{2}$, which becomes the final state.

For the final iteration, $i = 3$, say we obtain $c = 0.7$. So far, we have $m_1(\hat{A}_{33}) = 1$ and $m_2(\hat{A}_{23}) = -1$. We must obtain $m_3(\hat{A}_{13}) = 1$, independent of the value of c , to be consistent. Let's check that. Indeed, according to postulate 2, since $\hat{\sigma}_x \otimes \hat{\sigma}_x (|00\rangle + |11\rangle)/\sqrt{2} = 1 (|00\rangle + |11\rangle)/\sqrt{2}$, $m_3(\hat{A}_{13}) = 1$ for all allowed values of c . As a remark, it maybe emphasised that the $m_2(\hat{A}_{33}) = m_3(\hat{A}_{33})$ and $m_2(\hat{A}_{23}) = m_3(\hat{A}_{23})$, which essentially expresses compatibility of these observables, viz. measurement of \hat{A}_{13} doesn't affect the result one would obtain by measuring operators compatible to it (granted they have been measured once before).

IV. IMPLICATIONS

A. Locality

Definition 14. A theory/model/map is *local* iff two spatially distinct points can not be influenced instantaneously.

Proposition 15. *A prediction map m consistent with QM must be non-local.*

Proof. [Bell's] Formally it is clear, since a violation of Bell's inequality implies non-locality. Consequently, any consistent completion of QM, will be non-local. More insight can be gained, although some care is needed. The notion of compatible observables is that commuting observables don't disturb each other and can be simultaneously known. While the statement is not precisely stated, we only need to note that $m_{k_1}(\hat{C}) = m_{k_2}(\hat{B}_1)m_{k_3}(\hat{B}_2)$, for $(k_1, k_2, k_3) \in \{(1, 2, 3), (2, 1, 3), \dots\}$ where m_i are defined as before. In words, this means that to measure \hat{C} , one can instead first measure \hat{B}_1 and then measure \hat{B}_2 . A multiplication of the values obtained, can be taken to be the value a measurement of \hat{C} yields. This can be verified by measuring \hat{C} subsequently. The Bell scenario offers an interesting freedom, by restricting the form of \hat{B}_i . To evaluate $\langle \hat{B} \rangle$, one is required to evaluate $\langle \hat{B}_1 \hat{B}_2 \rangle$. From the compatibility argument, we need to find the average value of $m_3(\hat{B}_1 \hat{B}_2) = m_1(\hat{B}_1)m_2(\hat{B}_2)$, viz. $\langle m_3(\hat{a}_1 \otimes \hat{b}_1) \rangle = \langle m_1(\hat{a}_1 \otimes \hat{\mathbb{I}}) \rangle \langle m_2(\hat{\mathbb{I}} \otimes \hat{b}_1) \rangle$. So far, the statements were general. Now we assume that our theory is non-multiplicative (and non-contextual). A consistent theory (with QM), using these assumptions, can and must violate $\langle \hat{B} \rangle \leq 2$. However, if the particles are taken far away, then the only way the theory can be non-multiplicative ($m_1(\hat{B}_1) \neq m_2(\hat{B}_1)$ for example), is if the theory is non-local. Locality will entail that $m_1(\hat{B}_1) = m_2(\hat{B}_1)$ for instance, since the information about which observable was measured first, can't be propagated instantaneously. We see therefore that non-locality arises quite naturally in consistent non-multiplicative theories. \square

Proposition 16. *[can perhaps be skipped] Consider a product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, corresponding to two particles. \mathcal{H}_A and \mathcal{H}_B are discrete (correspond to spins say). Let mutually commuting observables \hat{B}_i be s.t. $\hat{B}_i = \hat{\mathbb{I}} \otimes \hat{A}_{Bi}$ or $\hat{B}_i = \hat{A}_{Ai} \otimes \hat{\mathbb{I}}$, where $A_{Ai} \in \hat{\mathcal{H}}_A$ and $A_{Bi} \in \hat{\mathcal{H}}_B$ are arbitrary observables. If locality is assumed, then a prediction map m and it's associated sequence map s must be s.t. over these observables, sequential multiplicativity and multiplicativity become identical.*

B. Apparent Contextuality

We have already constructed an explicit non-contextual model, which is consistent with QM. This model we knew had to be non-multiplicative. We will see how non-multiplicativity gives rise to what one might confuse to mean contextuality. Imagine

$$\hat{B}_1 = \hat{\sigma}_z \otimes \hat{\mathbb{I}} = |00\rangle\langle 00| + |01\rangle\langle 01| - [|10\rangle\langle 10| + |11\rangle\langle 11|],$$

$$\hat{B}_2 = \hat{\mathbb{I}} \otimes \hat{\sigma}_z = |10\rangle\langle 10| + |11\rangle\langle 11| - [|00\rangle\langle 00| + |01\rangle\langle 01|],$$

while we define

$$\hat{C} = f(\{\hat{B}_i\}) = 0. |00\rangle\langle 00| + 1. |01\rangle\langle 01| + 2. |10\rangle\langle 10| + 3. |11\rangle\langle 11|.$$

\hat{C} maybe viewed as a function of \hat{B}_1 , \hat{B}_2 and other operators \hat{B}_i which are constructed to obtain a maximally commuting set. A measurement of \hat{C} , will collapse the state into one of the states which are simultaneous eignkets of B_1 and B_2 . Consequently, from the observed value of \hat{C} , one can deduce the values of \hat{B}_1 and \hat{B}_2 . Now consider $\sqrt{2}|\chi\rangle = |10\rangle + |01\rangle$, for which $m_1(\hat{B}_1) = 1$, and $m_1(\hat{B}_2) = 1$, using the C-angle model, with $c < 0.5$. However, $m_1(\hat{C}) = 1$, from which one can deduce that B_1 was $+1$, while B_2 was -1 . This property itself, one may be tempted call contextuality, viz. the value of B_1 depends on whether it is measured alone or with the remaining $\{B_i\}$. However, it must be noted that B_1 has a well defined value, and so does \hat{C} . Thus by our accepted definition, there's no contextuality. It is just that $m_1(\hat{C}) \neq f(m_1(\hat{B}_1), m_1(\hat{B}_2), \dots)$, viz. the theory is non-multiplicative. Note that after measuring \hat{C} however, $m_2(\hat{B}_1) = +1$ and $m_2(\hat{B}_2) = -1$ (for any value of c) consistent with those deduced by measuring \hat{C} .

C. Entanglement and Superposition

It is known that for simultaneous eigenstates of \hat{B}_i , multiplicativity must hold. Consequently, any violation of multiplicativity must arise from states that are superpositions (of the simultaneous eigenkets). Note however, that entanglement is not necessary to show a violation, since, for example, the PM test is a state independent test, where a separable state can be used to arrive at a contradiction. For demonstrating non-locality using Bell's proof, however, it can be shown that entanglement is necessary.

TODO: Complete this part.

-
- [1] A. Einstein, B. Podolsky, and N. Rosen, [Phys. Rev. **47**, 777 \(1935\)](#).
 - [2] S. Kochen and E. P. Specker, *Journal of Mathematics and Mechanics* **17**, 59 (1967).
 - [3] A. Gleason, *Indiana Univ. Math. J.* **6**, 885 (1957).
 - [4] J. S. Bell, [Rev. Mod. Phys. **38**, 447 \(1966\)](#).
 - [5] A. Peres, *Journal of Physics A: Mathematical and General* **24**, L175 (1991).
 - [6] N. D. Mermin, [Phys. Rev. Lett. **65**, 3373 \(1990\)](#).
 - [7] D. Bohm, [Phys. Rev. **85**, 166 \(1952\)](#).
 - [8] D. Bohm, [Phys. Rev. **85**, 180 \(1952\)](#).
 - [9] B. R. La Cour, [Phys. Rev. A **79**, 012102 \(2009\)](#).
 - [10] D. Dürr and S. Teufel, *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory* (Springer, 2010).