Assumptions of metric variable-type in Bell’s theorem

Fosco Ruzzene

Department of Econometrics & Business Statistics, Monash University 900 Dandenong Road, Caulfield East, Australia

Email address: Fosco.Ruzzene@monash.edu

To cite this article: Fosco Ruzzene. Assumptions of Metric Variable-Type in Bell’s Theorem. American Journal of Modern Physics. Vol. 2, No. 6, 2013, pp. 350-356. doi: 10.11648/j.ajmp.20130206.22

Abstract: An analysis of both the original and the CHSH Bell inequalities is presented. Two additional mathematical assumptions are identified in the theorem. These are: all variables in the inequalities have a field algebraic structure, and all variables have measurability as a mathematical property. This means the variables are of metric-type, mathematically indistinguishable from those of classical theories. The consequences of attributing the violation of Bell’s theorem to these assumptions are examined.

Keywords: Bell’s Inequalities, Non-Classical Measurability, Metric Variables, Projective Geometry

1. Introduction

Bell’s theorem is considered to be the testing ground on fundamental questions on the nature of objective reality [1]. Far reaching conclusions have been drawn on the theorem’s failure to predict experimental outcomes.

By mathematically modelling an ensemble of correlated particle pairs, the theorem obtains inequalities which contradict the corresponding quantum analysis. Bell’s theorem is not however primarily about reproducing experiment but rather testing underlying assumptions. Traditionally, two assumptions, usually termed “realism” and “locality”, have been identified. Violation of the inequality, which has been reported in many experiments, notably A. Aspect et al (1982) [2], leads to the conclusion that “one or other or both” of these classical notions of reality must be rejected.

Clearly such logic is correct only if no further assumptions are present. If additional assumptions are being made the implications of the theorem’s violation would be more complex. It is this possibility which will be the focus of this discussion. Questions of validity or primacy of identified assumptions, while in no way diminished in significance, will be of lesser relevance here.

Although it was non-locality which captured Bell’s interest and motivation, “realism” has also attracted theoretical interest [3]. Leggett developed new inequalities which abandon locality but maintain realism and which are also in disagreement with QM. Subsequently, experimental violation was confirmed [4]. However, notable contributors, while stressing the importance of this line of investigation, argue that due to the restrictive nature of Leggett’s theory, rejection of “realism” is not a matter of logical deduction [5,6].

Nevertheless, the consensus view, not to suggest that such matters be resolved by polling, is that locality is the critical inoperative assumption. Indeed, it is argued that since Bell’s inequality can be obtained with only the locality assumption, the “realism” assumption becomes secondary [7,8].

This sketchy outline is presented only for context. There is wide ranging discussion on the precise meaning of these concepts and their philosophical implications. These profound questions will not, however, be considered. Primarily, the focus of this discussion is the mathematical structure of the theorem and its underlying mathematical assumptions, particularly assumptions made about the properties of the variables involved in the inequalities. This work is not an attempt to solve the EPR paradox, introduce new thought experiments or extend the theorem with new mathematical formulations.

Experimental violation of the theorem leads some to draw conclusions on the nature of deeper hidden variables theories. Such thinking is motivated by the ERP questioning of whether quantum mechanics is complete. These important questions will likewise not be considered. This discussion is restricted to the more modest view that Bell’s theorem servers to focus more sharply the difference between classical and quantum theories.

Bell’s theorem has its origins in the EPR paradox. This celebrated construction introduces notions (EPR terminology will be used throughout) of “objective reality”,
which is independent of the “physical theory” and the “physical concepts” with which the theory operates; “physical concepts” are intended to correspond with “objective reality”. Also introduced is the idea of an “element of physical reality” which is to have a counter-part in the “physical theory”.

For the sake of definition, the term “variable” will refer specifically to an entity in a mathematical formulation having mathematical properties which somehow have correspondence with the properties of physical concepts. As will become clear, the translation from physical concept to variable is not a simple procedure.

While Bell’s theorem has its motivation in EPR it is fundamentally different. EPR does an operational analysis of a hypothetical QM experiment. Bell’s theorem is, however, a mathematical modelling of the proposed experiment. This raises its own set of problems. The most basic is that of authentically interpreting physical concepts into mathematical variables.

The role geometry plays in a physical theory is well understood by considering the classical theories. The structure of space and time is represented by the axioms of the underlying geometries. Both classical mechanics and relativity adopt a metrical geometry. As will be seen, geometry is also an influencing consideration in Bell’s theorem.

It is important to be clear what Bell’s theorem is not: it is not a statistical analysis of the experimental data where underlying structures and relationships between variables are extracted by the analysis. Rather, the theorem is a mathematical model involving algebraic symbols for variables supposedly representing physical concepts; where the variables are subject to imposed mathematical conditions and subject to appropriate rules of algebra and arithmetic. Although restricted to a limited number of specific experiments, in this sense, Bell’s theorem is a physical theory as par the EPR definition.

While Bell’s theorem attracts detailed analysis of the inequalities, both experimentally and theoretically, its broader context cannot be ignored. Bell’s primary motivation was to examine foundational questions of QM. Neither EPR nor Bell’s theorem are ends in themselves. These constructions are designed to critically examine basic concepts of QM. Bell’s theorem is central to the current lively debate on foundations question [6]. Such discussions have been seen as mainly philosophical in character and not technically required. Recent weak measurement experiments, probing the micro details of quantum particle movements, are seen as becoming powerful tools for investigating fundamental questions, which may challenge the orthodox view [9].

Weak measurement experiments have demonstrated that the uncertainty relation formulated by an operational analysis of the Heisenberg microscope is incorrect [10,11]. While orthodox QM re-formulated, the hypothetical experiment was central to its initial development. Trajectories of quantum particles moving through interferometers have also been experimentally tracked [12,13,14]. These developments give at least some empirical justification for seriously considering theories based on what has been termed a “primitive ontology” [15].

2. Bell Inequality

An overview of the Bell inequality and corresponding experiment is useful. Briefly, an operational description of the experimental arrangement to test Bell’s theorem in the original form can be summarised as follows: a pair of half spin particles in the singlet state move in opposite directions [16]. At location A, measurement of spin component σ₁ a along unit vector a is made for particle 1. Similarly, at location B, σ₂ b is measured for particle 2 in direction b. The results of both measurements are

\[ A(a, \lambda) = ±1 \quad B(b, \lambda) = ±1 \quad (1) \]

where \( \lambda \) is the hidden variable parameter which plays no part in this discussion. A requirement of QM is that if measurement of particle 1 spin yields the value +1 then the measured spin component of the other particle must yield −1, and vice versa. The expectation value of the product of the two components is given by the expression

\[ P(a, b) = \int d\lambda \rho(\lambda)A(a, \lambda)B(b, \lambda) \quad (2) \]

Multiplying the measurement variables together is the mathematical condition to incorporate Locality. Bell introduces the critical condition that

\[ A(b, \lambda) = - B(b, \lambda) \quad (3) \]

Equation (2) becomes

\[ P(a, b) = - \int d\lambda \rho(\lambda)A(a, \lambda)B(b, \lambda) \quad (4) \]

By introducing a third unit vector in direction c, together with some simple algebra subject to the field properties of real numbers the inequality results as

\[ 1 + P(b, c) \geq |P(a, b) - P(a, c)| \quad (5) \]

Bell showed that this inequality contradicts the predictions of QM.

Relation (3) is critical: the RHS symbols, \( B(b, \lambda) \) is the variable representing the outcome of the actual physical measurement of the spin of particle 2 at location B in direction b; \( A(b, \lambda) \) is the purely mathematical variable representing the spin value of particle 1 in direction b; critically, experimentally, this is not a measured value. Yet mathematically the two variables, according to (3), must have the same properties. Bell assigns number-values to both variables simultaneously. This meets the ERP element of reality criteria on the spin value of particle 1 along the unmeasured b direction.

Relation (3) introduces two fundamental mathematical assumptions. Firstly, both variables, since they take number-values, have the algebraic structure of a field.
Secondly, both variables possess the mathematical property of measurability. Both variables are metric-type variables of classical physical theories.

It is the second assumption which is of more immediate interest. Examining whether a mathematical variable somehow counter-parts measurability, or lack of it, may appear to be a novelty. However, this is not so. Recent experiments on fundamental aspects of the uncertainty principle have established that the measurement-apparatus plays no fundamental role in QM [11]. Non-classical measurement features of QM must then be attributed to its physical concepts. Hence mathematical variables in any physical theory which attempt to describe quantum objective reality must in some fashion incorporate such non-classical measurability.

Alternatively, if the ERP element of physical reality criteria requires that a representing variable have the mathematical property of being able to take number-values why not also have a characteristic mathematical feature to incorporate non-measurability? Quantum mechanics does so by requiring that the variables which express the theory have a ring rather than field algebraic structure. As will be shown shortly, variables in the classical theories have the dual properties of being able to be assigned number-values and also measurability. While these attributes of metric-type variables may seem the “natural order” that is not the mathematical reason why classical variables have such properties.

To illustrate further the assumption of measurability of the Bell variables consider a purely hypothetical modification to the experiment. Suppose a “classical” Stern-Gerlach apparatus at A which measures both spin components $\sigma_a$ and $\sigma_b$ of particle 1 simultaneously, and likewise at location B. Clearly this is a purely hypothetical experiment which is completely unphysical!

The expectation value of the product of the two components for the same particle is

$$P_i(a, b) = \int d\lambda \; p(\lambda)A(a, \lambda)A(b, \lambda)$$  

(6)

where $A(a, \lambda)$, $A(b, \lambda)$ are both outcomes of measured variables in the hypothetical experiment for particle 1. However, the integrant in (6) is the same as that of (4). Hence, the Bell inequality fails to mathematically distinguish between the real and unphysical experiment.

To avoid the perfect anti-correlation aspects of condition (3) the CHSH inequality has been developed for a modified experiment involving a system of four observables: $A$, $A'$ for particle 1 and B, $B'$ for the second particle. For the purpose of this discussion the derivation of Angel G. Valdenebro (2002) [17] will be followed. Possible values are denoted by the sets $\{a_i\}, \{a'_i\}, \{b_i\}$ and $\{b'_i\}$ which are bounded by ±1. Using the boundary constraint it can be shown that

$$\left| a_i(b'_i - b_k) + a'_i(b'_k + b_k) \right| \leq 2$$  

(7)

Mathematically all four values variables have the same two properties: all have a field algebraic structure, and either all have the property of measurability or none do so. Clearly the latter option cannot represent the experimental situation, and so is ignored. Condition (7) therefore applies also to the unphysical experiment with the “classical” SG apparatus. To obtain a combination of the expected value of the product of two observables a joint probability distribution function is introduced such that

$$\left| \Sigma_{i,j} p(a_i, a'_j, b_k, b'_l) [a_i(b'_j - b_k) + a'_j(b'_l + b_k)] \right| \leq 2$$  

(8)

Predominately from the properties of the probability function, Valdenebro examines the usual range of assumptions believe to underpin Bell’s theorem. The list includes: “realism”, which incorporates hidden variables (deterministic or stochastic) and reality as single valued; “fair distribution” which incorporates no selective efficiency, no conspiracy and no backward causation; and “Bell-Locali” incorporating parameter independence and outcome independence. The analysis does not identify a primal assumption.

Applying various conditions to the JPD, which apply also to the “classical” SG apparatus, including that of factorbility for the Locality criteria, gives the inequality

$$\left| <AB> + <A'B> + <A'B> - <AB> \right| \leq 2$$  

(9)

These various assumptions, and whether one can be identified as primal to the others, are interesting but not central to this discussion. The central condition here is condition (7), where the variables are assigned number values and the boundary conditions calculated by simple arithmetic. These basic operations are only valid under the assumption that the variables are metric-type.

The CHSH inequality has been generalised to what is claimed to define global constraints on any experimental situation not just quantum [18]. In this refinement, at location 1 the settings on the measurement-apparatus are designated by $a$, $a'$ and at location 2 by $b$, $b'$. Under the standard assumptions an inequality, equivalent in form to (9), is obtained for a combination of expectation values $E(a, b)$, $E(a, b')$, $E(a', b)$ and $E(a', b')$. These are now functions of the apparatus settings; hence the claim of generality. Operationally, the experiment is performed to the requirements of QM where at each location only one of the two possible settings at selected at random. However, as previously, there is no mathematical condition in the derivation of the inequality prohibiting simultaneous measurement at the one location. Analysis of the refined experimental with a “classical” SG apparatus at each location gives an inequality which is mathematically indistinguishable from the that with the quantum SG apparatus. As with the original Bell inequality, this inequality also fails to distinguish between the real and unphysical experiments.

It should be emphasised that the introduction of the hypothetical experiment is purely illustrative: the critical issue is that with the inequalities all variables have the same mathematical properties even though they are supposed to represent quantities which have different physical
properties.

Attributing the violation of Bell’s theorem to the inoperativeness of these additional assumptions leads to the conclusion that the mathematical variables of the physical theory describing quantum objective reality cannot be the metric-type variables of classical theories. But this is already known!

The EPR paradox is based on two basic premises. Firstly, it is accepted from QM that if the operators corresponding to two physical quantities do not commute “then the precise knowledge of one of them precludes such knowledge of the other”. And secondly, the condition that both have simultaneous reality requires the both have definite values.

Bell “interprets” these statements of physical concepts into a mathematical model. The most basic requirement is then to construct appropriate variables with the appropriate mathematical properties. Bell’s metric-type variables, since they can take number-values, obviously fulfil the second condition.

However, to the extent that “knowledge” in the EPR definition can be interpreted to mean measurability for a mathematical variable, Bell’s metric-type variables do not adequately represent EPR’s first condition. Since all variables in Bell’s model have the mathematical property of measurability, which is a consequence of condition (3), there is no mathematical characteristic in the theorem where “knowledge of one precludes knowledge of the other”.

From this discussion it would seem that: Bells’ theorem fails to distinguish between real and unphysical experiments, says nothing new about the mathematical nature of quantum variables, and does not fulfil the conditions of the EPR paradox!

This does not mean to imply that the theorem is useless: its motivation is only to test underlying assumptions. Rather, Bell’s theorem, at least with this discussion, draws attention to the nature of quantum variables. Originally, the peculiar ring structure of QM variables was introduced for the pragmatic reason that the non-commuting algebra fitted experiments. With subsequent formulations of QM, representing observables by operators in Hilbert space is an axiom. Consequently, the origins of quantum variables, compared to corresponding to classical variables, lack clarity.

To further pursue this line of reasoning, it is worth clarifying how metric-type variables come about in classical theories. Classical mechanics is based on the structure of space and time described by Euclidean geometry. It should be stressed that space and time as aspects of objective reality do not underpin the classical theory. Rather, it is the axioms of Euclidean geometry which fulfil this mathematical role.

With Euclidean geometry its axioms permit purely geometrical constructions of what Hilbert calls the algebra of segments (referred to as the algebra of points in other treatments on foundations of geometry) [19]. These are geometrical rules for combinations of pairs of segments to give a unique third segment. As Hilbert shows, these geometrical constructions form a field. Real numbers under the rules of basic arithmetic also form a field. Hence, due to the isomorphism between the algebra of segments and real numbers, segments (and also points) in Euclidean geometry can be represented by real numbers.

The axioms of Euclidean geometry also permit comparison of distances between points which is the basis of measurability. Hence measurability is also a property of variables in a physical theory underpinned by Euclidean geometry. While these features are true of all metrical geometries, including those on which relativity theories are based, they are not universal. Projective geometry does not sustain measurability, and in general does not define an isomorphism between points and real numbers via a field algebraic structure.

The purpose of this brief refresher tutorial on foundations of geometry is to remind that the properties of variables in a mathematical model must be justified mathematically. There are fundamental reasons why the variables of classical theories have the metric properties that they have. These deeper considerations are absent in Bell’s theorem. Metric properties are simply assumed: even for spin the supposedly least “classical” of quantum quantities.

A Bell inequality has also been obtained for position, momentum and energy variables for particles in a potential [20]. This inequality, which is also in disagreement with the predictions of QM, is considered to be closer to the original EPR experiment. In obtaining the inequality, the space and time variables are assumed to form an algebraic field. As in the original Bell inequality, this assumption is necessary to fulfil the EPR element of reality criteria.

While no statement of geometry is made, since the variables involve space and time an underlying geometry must also be assumed. Assuming the variables form an algebraic field requires assuming also that the underlying geometry is simultaneously for both position and momentum Euclidean. Consequently, the variables in the inequality must possess the mathematical property of simultaneously measurability. Hence, non-simultaneous measurability is again not incorporated as a mathematical property of this inequality.

Although receiving little attention, Bell also analysed (said to be one of his last papers) the original EPR experiment for free particles with position and momentum variables [21]. Surprisingly, the results show no disagreement with the predictions of QM. The analysis is based on the Wigner phase-space distribution with the EPR wave function expressed as a delta function. A subsequent treatment, verifying Bell’s original conclusion, used Gaussian functions [22]. Both works also give a CHSH-type inequality which, consistent with the Wigner analysis, shows no violation.

Bell’s analysis again focuses on non-locality, and in this case, contradiction with the completeness of the EPR wave function. It is worth quoting Bell’s own conclusion: “with the wave function (original EPR), then, there is no non-locality problem when the incompleteness of the wave function description is admitted”. This would seem a rather
startling conclusion.

This discussion, however, is more focused on the mathematical properties of the variables involved in the Wigner function. Since for free particles momentum is a constant of motion, the momentum eigenvalue is a metric-type variable. Hence, both position and momentum variables in the Wigner function are of metric type. The assumption of using metric-type variables is then valid.

A different analysis to that of Bell, also for free particles, has been reported where position measurements are made on the same particle at different times [23]. Since position operators at different times do not in general commute a CHSH violation is expected. The analysis is done for three different conditions, for one of which, the operators are commutative. While violation of the CHSH inequality is found for the conditions of non-commutative operators, no violation is found with the commutative case.

For the commutative case the variables form a field, the assumption of metric-type variables is then valid. This is not so for the non-commutative case where variables form a ring.

While it can be legitimately argued that Bell inequalities are violated in some cases and not others, and that definitive conclusions should not be drawn from a restricted family of cases, nevertheless, these free particle treatments do provide at least some experimental connection to the issues raised in this discussion. Although restricted to just four experimental arrangements there is a small pattern: the situations where the non-violation free particle results. The possibility has been explored that QM can be understood from a ring structure. An isomorphism then exits between such variables. This raises philosophical objections to represent the structure of space by Projective geometry. However, since under special case conditions Euclidean geometry is obtained from Projective geometry, subsets of points have the usual Euclidean properties. All measurability features of QM, in particular expectation and eigenvalues which play a significant part in the free particle experiments, are reproduced.

While addressing the “profound questions” on the nature of reality is not within the scope of this discussion, some reference to “realism” is appropriate. Factorability plays no part in the assumptions identified here. There are two immediate aspects to the “realism” question: is there a measurement-independent reality? And, how is measurability represented mathematically? It is worthwhile considering the second question; the first is too profound.

EPR requires that measurability is represented mathematically by assigning number-values. However, this definition may be too restrictive. Numbers are, despite their awesome powers, just mathematical constructions subject to certain rules of algebra, as are various other entities.

It is worthwhile considering the most basic physical quantity: the distance between two points. A fundamentalist application of EPR would require that the distance is given by “x units of length”, where x is a number. Should such a statement not be possible then “distance” is not a physical concept and cannot be included in a physical theory. However, distance is itself generically defined from relationships between points which are determined geometrically. This raises the question whether constructions of sets of points can also be considered as identifying realism; especially, constructions which have invariant properties.

Since distance is not invariant with Projective geometry its introduction must begin from generic definitions. Concerning any two points P and Q the distance between these points, defined as (PQ), has certain fundamental properties:

\[
(PP) = 0 \quad (10a)
\]

\[
(PQ) = -(QP) \quad (10b)
\]

\[
(PQ) + (QR) + (RP) = 0 \quad (10c)
\]

for P, Q, R collinear points. Based on these properties the projective distance between P and Q is defined by the expression

\[
(PQ) = k \ln\{PQ, OU\} \quad (11)
\]

where k is a scale constant, O and U are chosen fixed points on a line, and \{PQ, OU\} is the cross ratio. It is easily shown that this definition fulfills the three generic properties.
Furthermore, under special case conditions Euclidean definitions can be obtained.

Adopting a broader definition of measurability to include relationships between points would overcome philosophical objections to describing the structure of space with Projective geometry. Clearly, this requires some departure from classical ideas. However, unless further assumptions are identified in Bell’s theorem showing otherwise, violation of the inequalities calls for rejection of some classical notions of objective reality.

Variables associated with Bell’s theorem are dichotomously defined as non-contextual (measurement reveals a pre-existing inherent property) and contextual (measurement dependent). Inequalities can be equivalently obtained by the assumption of NCHV alone, and by that of locality alone, although the latter is claimed to the primal assumption [7].

An inconsistency arises for NCHV for the bound and free particle experiments. For the same particle, for example, an electron in the ground state of hydrogen which escapes the Coulomb potential, an inherent property cannot exist and then cease to exist. The CHV definition is also faced with experimental complications. Weak measurement experiments showing the Heisenberg microscope analysis to be incorrect places a direct challenge to the Bohr construction of measurement-dependence. If the measurement-apparatus does not play a fundamental role it then becomes difficult to define measurement-dependent variables.

By extension of word formulation definitions of both variable types could be modified to accommodate any possible experimental inconsistencies. It would be desirable, however, to have definitions which are not constructed solely from experimental outcomes.

Nevertheless, experimental evidence does require, in some fashion, rejection of inherent properties; although the situation becomes complicated if Bell’s conclusion for the free particle case is not ignored. Likewise, experiments do suggest that some kind of contextual aspects are required.

Variables defined by the geometry proposed, while reducing measurement from fundamental to a geometrical-mathematical property, both retain inherent characteristics as well as having contextual-like properties. Their geometrical-mathematical properties, being assigned number-values and being measurable, are only possible under geometrically defined conditions. There is no dichotomy in this definition. Furthermore, the variable properties are not tailored from quantum experiments.

Locality is defined by concepts of space-separation, and time-separation for communication, between locations. These are metric concepts. For the particle whose path is described by a non-metric geometry this definition would require re-defining.

3. Conclusion

Apart from the speculative possibility, this discussion is solely an analysis of additional assumptions present in Bell’s theorem. This is not an attempt to somehow invalidate or disprove the theorem. What emerges is that there are additional assumptions giving rise to additional consequences which allow for an alternative line of investigation from the much considered questions of realism and non-locality.

References

[16] J. S. Bell, Physics, 1, 195 – 200 (1964)