Comments of Stephen Parrott
on
"A derivation of quantum theory from physical requirements"
by
Lluís Masanes and Marcus P. Müller

November 24, 2012, revised December 5, 2012

1 Addition December 6, 2012

The most important of the following comments points out what I suspected might be an essential error in the proof of Theorem 2 of the commented paper of the title. A draft had been sent to the authors. After no substantive replies had been received, I posted it on my website, www.math.umb.edu/ sp. The circumstances are more fully described in Section 5.

Since then, the authors have replied and convinced me that the conclusion of Theorem 2 is correct. Naturally, I want to correct any false impressions that the posted version may have created. However, I still think that the published proof of Theorem 2 may puzzle many readers, and that it would be a service to make public the explanation that one of the authors kindly provided me.

I have decided that the most efficient way of doing this would be to add an appendix to the original version. I do not intend to alter the original version, except that I will correct any typos that I find. The original only referred to a "possible error" in the proof of Theorem 2, and I think that is still valid in that it does point out what seems to me a gap in the proof.

2 Introduction

Over a year ago, I became interested in recent attempts to "derive" quantum mechanics from axioms which are physically intuitive. This is generally known as an "operational" approach. Some of the papers which I read, or attempted to read to various degrees, were:

Hardy, L., Quantum Theory From Five Reasonable Axioms, arXiv:0101012v4

Barrett, J., Information procession in generalized probabilistic theories, Phys. Rev. A **75** 032304 (2007), arXiv:0508211v3

Masanes, L., and Müller, M., A derivation of quantum theory from physical requirements, New J. Phys. 13:063001 (2011), arXiv:1104.1483v4, called MM below

Chiribella, G., D'Ariano, G., and Perinotti, P., Probabilistic theories with purification", Phys. Rev. A. 81, 062348 (2010), arXiv:

arXiv:0908.1583, (called CDP10 below)

Chiribella, G., D'Ariano, G., and Perinotti, P., Informational derivation of quantum theory, Phys. Rev. A 84,012311 (2011), arXiv: arXiv:1011.6451 (called CDP11 below)

Some remarks about my attempts can be found in the July 1, 2012 entry on the "papers" page of my website, www.math.umb.edu/∼sp . In summary, I wasn't able to penetrate any of these to the degree necessary to convince myself of their soundness.

I put the most effort into Hardy and the CDP papers. Hardy defines the physical situation unusually clearly and charmingly, and all of the papers mentioned above clearly owe a great debt to it. Their approaches are all similar, though not identical. However, I wasn't able to follow Hardy's mathematics in detail. It is not presented in a precise way, and many of the main results are relegated to appendices which I found too sketchy to verify in detail.

The CDP papers are extremely intricate. I found what seem to me to be errors, or at least questionable points. However, even if they turn out to be errors, the work is so intricate that I can't guess if they might be essential errors. A letter to Dr. Chiribella describing some of my questions can be found on the above web page. However, it is so technical that it will be incomprehensible to anyone not actively reading CDP in detail.

Barrett does not attempt to derive quantum mechanics, but formulates a framework similar to Hardy's in which such a derivation might be attempted. I have not read the full paper in detail, but have obtained useful insights from parts of it. I am not sure I can agree with all of its arguments.

Very recently (October, 2012), my attention was attracted to MM, and I began to read it in detail. In a way, it is my favorite of all the above because it is clearly written, and sufficiently simple compared to the above that I thought I might be able to convince myself of its key assertions in a reasonable time.

The authors of MM have posted an expository summary [6], which explicitly points out what its approach does *not* do; e.g., it does *not* address the so-called "measurement problem" of what happens to a state after measurement. This is refreshing in a field in which hype and self-promotion are so common. In the discussion to follow, I may treat MM and [6] as the same paper, in case it seems too distracting to distinguish between them.

That said, I must record my deep disappointment, and even sorrow, that I can't regard MM's derivation of quantum mechanics as valid, due to what seems to me a serious error in one key proof and inadequate justification of another. These will be discussed below.

Assuming that what look to me like errors actually are errors, they seem serious enough to require major rewriting. I do not know if they could be repaired, but I certainly would not discount the possiblility.

I am very concerned by the possibility that I could be mistaken about the errors. I have made mistakes before, not all that many, but enough to be acutely aware of my fallibility. I would greatly appreciate the opportunity to discuss MM in detail with any qualified person who is prepared to discuss it in detail. (The same goes for Hardy, CDP, and Barrett.)

I am trying to approach this exercise in the spirit of: "this is an interesting and promising approach; let us think about how it might be completed". I would love to see a convincing way to obtain the mysterious structure of quantum mechanics from simple physical principles. I hope the reader will approach it in the same spirit.

3 Possible errors

I have debated if I should present first what look like the most crippling errors, or go through the paper in the order that it was written. The latter would risk distracting the reader with what might be minor errors. I have decided to use a mixed approach, going through the paper as it was written, but not necessarily initially mentioning every little questionable point.

I think there may be questionable points in MM's setup, but I haven't noticed any important errors. A reader who is already familiar with MM may prefer to skip the following subsection on MM's basic setup.

3.1 MM's setup

Any undefined notation will be as in the paper. I hope I can reproduce most of the paper's notation. One exception is that my version of LaTeXdoes not include a "blackboard bold" font, used to denote Euclidean space R^n, as in MM's equation (3). When it is necessary to distinguish, for example, MM's equation (3) from my equation (3), I will use MM(3) for the former.

1. MM(1) presents MM's notation for states:

$$\psi = \begin{bmatrix} 1 \\ p(x_1) \\ \cdot \\ \cdot \\ \cdot \\ p(x_d) \end{bmatrix} = \begin{bmatrix} \psi_0 \\ \psi_1 \\ \cdot \\ \cdot \\ \cdot \\ \psi_d \end{bmatrix} \in \mathcal{S} \subset R^{d+1} \qquad \text{MM(1)}$$

MM does not fully explain why it inserts the redundant zero'th component 1. My guess is the following.

Call the vector $(\psi_1, \ldots, \psi_d)$ an "unaugmented" state, to distinguish it from MM(1). Hardy [1] and Barrett [2] use unaugmented states. For them, the procedure of making a measurement on a system in that state corresponds to a convex-linear map from the state space to the reals $R = R^1$. This is physically natural.

The trick of replacing the unaugmented state with MM(1) corresponds the convex-linear map just mentioned to a *linear* map from R^{d+1} to R, perhaps a more usual mathematical object. This is an example of what MM calls an "effect". One of MM's Requirements (axioms) is that *any* linear map is an "effect" (i.e., that any linear map corresponds to a physically possible measurement).

Evaluation of the zero'th component is now an "effect" which sends any state into 1. Later (in the proof of Theorem 1), this effect is named $\mathbf{1}(\cdot)$, a notation which does not seem to be explicitly defined.

In the formulation of Barrett, a particular measurement may have several possible results. For example, if a classical measurement is tossing a coin, there will be two possible results; if the measurement is throwing a cubical die, there will be six. The number of possible results for a measurement called x will depend on x, and will temporarily be denoted R_x, e.g., $R_{\text{coin}} = 2$ and $R_{\text{die}} = 6$.

To each measurement x and each possible result r of that measurement, $1 \leq r \leq R_x$, will correspond a probability $p(x, r)$. The notation greatly simplifies if we assume (unlike Barrett) that $R_x = 2$ for all x. In that case, all probabilities are completely determined by the probabilities $p(x, 1)$ because $p(x, 2) = 1 - p(x, 1)$, so we may simplify the notation by writing $p(x) := p(x, 1)$.

This is what MM seems to do, though it doesn't say so explicitly. More general cases in which some measurements have $n > 2$ results can be simulated by n successive two-outcome measurements (each on a newly prepared state).[1] Because MM is not explicit about this point, its definitions of "n-outcome measurements", "distinguishable states", and "capacity" on p. 3 seem logically ambiguous as stated. Where these were mentioned, I have made my best guess at the author's intended meaning.

Before leaving this topic, I should point out that Hardy [1] uses unaugmented states *without* assuming that $p(x, 2) = 1 - p(x, 1)$. He considers the possibility that the device which "presents" a state to the measuring apparatus sometimes fails, so that there is only a probability q that a valid (nonvoid) state impinges on the measuring device. In such a situation, we would have $p(x, 2) = q - p(x, 1)$, and Hardy generalizes the notion of "state" to consider this to be an "unnormalized state". (A "normalized" state would have $p(x, 2) = 1 - p(x, 1)$.) Hardy's treatment is quite vague regarding these new "unnormalized" states. Some key points of MM's argument cite statements of Hardy without noting that Hardy's basic setup differs significantly from MM's. I cannot always follow Hardy's reasoning, and even when I can, it is not always clear to me that Hardy's conclusions can be taken over verbatim to MM's context.

[1]For example, the first two-outcome measurement would have outcome "Yes" if the n-outcome neasurement with outcomes $r = 1, 2, \ldots, n$ resulted in $r = 1$, and "No" otherwise. If the first measurement is "Yes", stop; otherwise perform a second measurement on a newly prepared state to see if $r = 2$, etc. Then the probability of $r = 2$, say, is defined to be the observed conditional probability $p($ second measurement "Yes" | first measurement "No" $)$.

3.2 Theorem 1

"**Theorem 1.** A state in $\hat{\mathcal{S}}_2$ is pure if and only if it belongs to the boundary $\delta\hat{\mathcal{S}}_2$."

The nontrivial part is the "if". The circumflexes ("hats") refer to the Bloch representation which has just been defined in equation MM(13). However, the Bloch representation does not play any role in the proof of Theorem 1.

Before delving into the proof, let us think about MM's Requirement 3 for the special case of a system $\mathcal{S}_2$ with capacity 2, which will be the case relevant to the Theorem. Requirement 3 states that if Ω_1, Ω_2 is a complete measurement on $\mathcal{S}_2$, then the set of states $\psi \in \mathcal{S}_2$ with $\Omega_2(\psi) = 0$ is equivalent to $\mathcal{S}_1$. It follows from MM's wording of Requirement 3 that *all systems of capacity 1 are equivalent*, where MM defines "equivalent" as "related by an affine transformation".

It is natural to ask for an explicit description of $\mathcal{S}_1$. The first paragraph of the proof of Theorem 1 provides a very simple one (though it does not state it in this way): any realization of $\mathcal{S}_1$ consists of just one point.

The proof of Theorem 1 constructs a state $\hat{\phi}_{\text{one}}$ and a so-called "tight"[2] effect $\hat{\Omega}_{\text{one}}$ such that $\hat{\Omega}_{\text{one}}(\hat{\phi}_{\text{one}}) = 1$ and $\hat{\Omega}_{\text{one}}(\hat{\psi}) < 1$ for $\hat{\psi} \neq \hat{\phi}_{\text{one}}$. "The two effects $\hat{\Omega}_{\text{one}}$ and $\mathbf{1} - \hat{\Omega}_{\text{one}}$ define a complete measurement on $\mathcal{S}_2$". This implies that the one-point set $\{\hat{\phi}_{\text{one}}\}$ is a realization of $\mathcal{S}_1$.

The fact that any two realizations of $\mathcal{S}_1$ are *equivalent* in MM's technical sense does not imply that they are physically indistinguishable. For example, the two one-point state spaces

$$\left\{ \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} \right\} \quad \text{and} \quad \left\{ \begin{bmatrix} 1 \\ 1/3 \end{bmatrix} \right\}$$

are equivalent (because the only transformation between two one-point affine spaces is trivially affine), but most people would probably think of them as physically distinct.

This is not a criticism of MM, but a general feature of axiomatizations. One of the purposes of axiomatization is to strip away "irrelevant" aspects of a situation so as to concentrate on more important aspects. For example, the two-dimensional vector space S consisting of all real solutions of the differential equation $y''(x) + y(x) = 0$ can can be studied by proving theorems about the Euclidean plane R^2. Though the set of solutions has additional structure not posessed by R^2, for many purposes they can be considered as "equivalent", and it may be easier to think about R^2. In MM's context which attempts to "derive" quantum mechanics, it is a legitimate issue whether their axioms may strip away structure which should be considered relevant.

3.3 Possible error in the proof of Theorem 2

The essence of Theorem 2 is the claim that $\mathcal{S}_2$ is equivalent to a unit ball in R^{d_2}, for some d_2. MM's statement is:

[2] A "tight" effect Ω is defined as one for which there are two states ψ_0 and ψ_1 such that $\Omega(\psi_0) = 0$ and $\Omega(\psi_1) = 1$.

“ **Theorem 2.** *There is a set of fiducial measurements for which $\hat{\mathcal{S}}^2$ is a d_2-dimensional unit ball.*”

Here the circumflex on $\mathcal{S}^2$ refers to the Bloch representation. According to Theorem 2, in the Bloch representation, $\hat{\mathcal{S}}^2$ *is* a d_2-dimensional unit ball; the abstract $\mathcal{S}^2$ is *equivalent* to a unit ball.

It is not clear if the authors intend to imply additionally that d_2 is the “dimension” of $\mathcal{S}_2$, where “dimension” is defined as the “d” of MM(1) (which is the minimal number of probabilities necessary to specify the state). The latter does not follow from formal logic, but many readers might make this guess.

The statement of the theorem suggests that construction of the fiducial measurements will establish the equivalence of $\mathcal{S}^2$ with the unit ball, but the proof first attempts to establish the equivalence, and then routinely constructs the fiducial measurements.

It seems to me that the attempted proof of Theorem 2 does not actually show that $\mathcal{S}_2$ is equivalent to the unit ball. I think that it shows only that $\mathcal{S}_2$ is equivalent to a *subset* of the unit ball which is invariant under a *subgroup* of $O(d_2)$, where $O(d_2)$ denotes the orthogonal group for R^{d_2}. For an example of such a subset which is not the entire ball, consider a disc whose circle boundary is an equator of the unit sphere in R^3. This is invariant under the subgroup of rotations about an axis perpendicular to the disc.

I think that Theorem 2 implicitly assumes what it is trying to prove, namely that $\mathcal{S}_2$ is affinely equivalent to the unit ball in R^{d_2} for some d_2. First it transforms the “standard” representation of MM(1) into the “Bloch representation”. In the Bloch representation, the pure states of $\mathcal{S}_2$ all lie on a sphere in R^{d_2} centered at the origin. By a trivial normalization, we may assume this is the unit sphere. In the following, the “hats” on the various quantities refer to the Bloch representation, and the primes to the normalized version. This is MM’s notation. The proof may look less formidable if the reader ignores the hats and primes and just thinks of $\mathcal{S}_2$ as a subset of the unit ball of R^{d_2}

The proof now makes the crucial claim:

> “Since Theorem 1 also applies to the redefined set $\hat{\mathcal{S}}'_2$, it must be a unit ball.”

But Theorem 1 only states that every point of the boundary of $\mathcal{S}_2$ ($\hat{\mathcal{S}}'_2$ in the present notation), is pure (and conversely). The boundary of $\hat{\mathcal{S}}'_2$, is a *subset* of the unit sphere in R^{d_2}, but it has not been established that it consists of *all* of the unit sphere. In the present Bloch representation, we know that $\mathcal{G}_{d_2}$ consists of orthogonal linear transformations of R^{d_2}, but we do not know that it consists of *all* orthogonal linear transformations. (If it did, we could transform any pure state in the unit sphere to any other by some transformation in $\mathcal{G}_2$, so that $\hat{\mathcal{S}}'_2$ would contain the unit sphere and hence the unit ball.)

The rest of the proof of Theorem 2 routinely constructs the claimed fiducial effects *given* its conclusion quoted above that $\hat{\mathcal{S}}'_2$ is the entire unit ball. The statement of Theorem 2 may give the impression that the construction of these fiducial effects implies that $\hat{\mathcal{S}}'_2$ is a unit ball, but such an impression would be wrong. The construction of the fiducial effects is valid only if it is already known that $\hat{\mathcal{S}}_2'$ is the entire unit ball.

3.4 The role of Theorem 2 in the paper

For perspective, it may be useful to indicate the essential way in which Theorem 2 is used in the rest of the paper. MM defines a "Bloch representation" of $\mathcal{S}_2$ such that in the Bloch representation, $\mathcal{S}_2$ is a subset of R^{d_2} for which any transformation in the transformation group $\mathcal{G}_2$ for $\mathcal{S}_2$ is an orthogonal linear transformation. This is done by using the freedom of choosing an origin for an affine space to choose an origin which is fixed by all transformations in $\mathcal{G}_2$. In the Bloch representation, $\mathcal{G}_2$ is a subgroup of $O(d_2)$, the group of all othogonal linear transformations on d_2.

Working in this Bloch representation, if there is any pure state in the boundary of $\mathcal{S}_2$ (which there is from Theorem 1), applying all transformations in $\mathcal{G}_2$ to that pure state will produce all pure states by Requirement 4. By a trivial normalization, all pure states may be taken to have norm 1, i.e., $\mathcal{S}_2$ is a subset of the unit sphere of R^{d_2}.

If we knew that *all* points of the unit sphere corresponded to pure states of $\mathcal{S}_2$ (which is the conclusion of Theorem 2), we could hope to identify $\mathcal{G}_2$ by identifying all subgroups of $O(d_2)$ which are transitive on the unit sphere.

The latter problem has been analyzed in mathematical literature (with which I am not familiar), and MM quotes the following.

- For d_2 odd and unequal to 7, the only such subgroup which is connected is the special orthogonal group $SO(d_2)$ (the group of orthogonal transformations of determinant 1).
- For $d_2 = 7$, there is another possibility.
- For d_2 even, there are many possibilities.

Next MM rules out the cases of even d_2 by referring to an argument of Hardy, which is only cited, not reproduced in MM's context and notation. Hardy's axioms include one (his Axiom 2) which is not a Requirement of MM, and I am not sure how it would follow from MM's Requirements. There is no discussion of this in MM. Moreover, Hardy's argument leading to the conclusion which MM cites is not clear to me.

In summary, Theorem 2 seems central to MM's main conclusions. Even assuming Theorem 2, there seem other aspects of MM's argument which deserve close scrutiny and more extensive exposition than given in MM.

4 A possible gap

MM presents a series of lemmas which imply that capacities are multiplicative: if c_A denotes the capacity of system A, then

$$c_{AB} = c_A c_B. \tag{1}$$

This is an essential ingredient for Hardy's argument mentioned above, and it is one of Hardy's [1] axioms. I could follow most of the reasoning leading to (1), except that MM uses many times the assumption, that a tensor product $\psi_A \otimes \psi_B$ of pure states ψ_A, ψ_B, is itself pure. I don't see how to prove this from MM's Requirements, though I wouldn't be surprised if it could be proved.

Several authors [4, 5, 7] have noted that the purity of $\psi_A \otimes \psi_B$ would follow immediately by "partial tracing" it to obtain states of A and B, respectively. An affine "partial trace" operation does not seem to exist in MM's framework, but it would seem a natural Requirement to add. Certainly, a state ψ_{AB} for AB must correspond to *some* state for A (and for B), and this correspondence should be affine for the usual reason.

5 Circumstances as of November 24, 2012

I hope the above makes clear that although I admire the work and straightforward writing style of authors Masanes and Müller, I have been unable to convince myself of the soundness of the work as a whole due to what seems to me an error in the proof of Theorem 2. Readers will be curious what the authors think of this. I would like to know, too.

On November 12, 2012, a first draft was sent to the authors, which was acknowledged by Dr. Masanes in a brief message which included no substantive comments. Email to Dr. Müller did not go through,[3] and to date I have been unable to contact him. On November 19, a nearly final draft was sent to Dr. Masanes with a message that if he had any objections to it, I would delay making it public, but no response has been received to date, November 24. The current version differs only cosmetically from the draft just mentioned.

The authors are probably busy people. and the interval from Nov. 12 to today, Nov. 24, is less than two weeks, so lack of immediate response may be understandable. I am posting the "Comments" because of personal circumstances which may prevent me from dealing with this matter for an indefinite period.

Finally, I want to say that I would not be surprised to find similar errors in other attempted "derivations" of quantum mechanics from operational principles. I was able to penetrate MM only because it is exceptionally clearly and simply written compared to other attempts. I am convinced that most physics papers are published without serious scrutiny and probably never read in detail.

[3]It was not returned as undeliverable in the usual way that happens, for example, when an address is mistyped. Instead, a message came back that it had been "delayed" and need not be resent. For several days "delayed" messages were received, followed by a final message that the server had given up trying to deliver the message.

6 Appendix added Dec. 6, 2012: Resolution of the "possible error" in Theorem 2

This appendix sketches a resolution of the "possible error" in Theorem 2. The main idea was kindly communicated to my by Dr. Masanes. I fleshed out the details, and if there should be any error in it, the error will be mine.

The essence of Theorem 2 (and the only part which is used in the rest of the paper, so far as I know) is that any system $\mathcal{S}_2$ of capacity 2 is affinely equivalent to a unit ball in a real Euclidean space of dimension d_2, where d_2 is the affine dimension of $\mathcal{S}_2$. That is what we shall prove.

The published proof of Theorem 2 is written in the context of a particular realization of $\mathcal{S}_2$, which MM calls the "Bloch representation". MM notationally distinguishes the Bloch representation by circumflexes on all quantities, and primes are added to refer to a simple normalization of the Bloch represntation. In the following, we shall simplify the notation by omitting the circumflexes and primes, it being understood that the Bloch representation is assumed unless otherwise specified.

The relevant facts about the Bloch representation of $\mathcal{S}_2$ are as follows:

1. $\mathcal{S}_2$ is a convex subset of a Euclidean space R^{d_2}; i.e, a real vector space of dimension d_2 with a nondegenerate inner product.

2. Elements of the group $\mathcal{G}_2$ of $\mathcal{S}_2$ are represented by linear transformations on R^{d_2} which preserve the inner product, i.e., by orthogonal linear transformations.

3. The affine dimension of $\mathcal{S}_2$ is d_2.[4]

   Informally, item 3 states that what one would normally call the "dimension" of the set $\mathcal{S}_2$ is the same as the dimension of the ambient space R^{d_2} containing $\mathcal{S}_2$. This rules out examples like a state space which is a two-dimensional disc in R^3.

Item 3 follows from the construction of the Bloch representation as affinely equivalent to the "standard representation" of MM(1), which has affine dimension the "d" of MM(1).[5] By construction, the Bloch representation "lives" in R^d, or R^{d_2} for the case of a system of capacity 2.

[4]The *affine dimension* of a convex subset C of an affine space is defined as follows. Choose any point $O \in C$ as the "origin", thus converting the affine space into a vector space. The affine dimension of C is then defined as the dimension of the subspace spanned by all elements of C. (This dimension is routinely shown to be independent of the choice of O.) This definition makes "affine dimension correspond to what one would normally call "dimension", e.g., a line segment has dimension 1, a disc dimension 2, etc.

[5]The identification of d as the affine dimension of the "standard representation" may be logically trickier than it might appear. To appreciate the potential logical problems, consider a one-point state space like $\{(1, 1/2)^T\}$. This, like all one-point affine spaces, has affine dimension 0, but it might appear that $d = 1$. On the other hand, if d is defined to be the minimal number of measurements necessary to specify the state, then it could be argued that $d = 0$ because there is only one state, and that the state should be represented as $\{(1)\}$. However, it seems unclear how one could know that there was only one state without making

Applied to a system $\mathcal{S}_2$, Theorem 1 asserts that its pure states consist of precisely its boundary points. When $\mathcal{S}_2$ is realized as a subset of R^{d_2} via the Bloch representation, the assumption (Requirement 4) that its group $\mathcal{G}_2$ acts *transitively* on pure states as *orthogonal* transformations implies that *all pure states have the same nonzero norm.* After a trivial normalization, we may take this norm to be 1. Thus the set of pure states is a subset of the unit sphere in R^{d_2}. We want to show that it is the entire unit sphere. Convexity will then imply that $\mathcal{S}_2$ is the entire unit ball.

To prove by contradiction that any point of the unit sphere is in $\mathcal{S}_2$. suppose that some point s of the unit sphere is *not* in $\mathcal{S}_2$. Next note that the origin 0 of R^{d_2} (corresponding to the maximally-mixed state of MM(11)) *is* in $\mathcal{S}_2$. Consider the largest $\lambda \geq 0$ such that λs is in $\mathcal{S}_2$. (MM assumes $\mathcal{S}_2$ to be compact, so there is such a λ.) Then $\lambda < 1$ because s is not in $\mathcal{S}_2$, and λs is a boundary point of $\mathcal{S}_2$ with norm $|\lambda s| < 1$. By Theorem 1, every boundary point of $\mathcal{S}_2$ is pure, so $|\lambda s| < 1$ contradicts the fact noted above that all pure states have norm 1. This contradiction shows that the assumption that there exists a point of the unit sphere which is not in $\mathcal{S}_2$ is untenable; hence $\mathcal{S}_2$ contains the whole unit sphere (and unit ball by convexity).

7 Appendix 2 added December 9, 2012

Just today, I received an extensive letter from Dr. Müller addressing various issues raised in the main "Comment". I am grateful for these clarifications.

Most of them involve semi-philosophical matters such as definitions and physical interpretations, but one points out an unjustified criticism at the end of section 3.4 which I would like to correct:

> "Next MM rules out the cases of even d_2 by referring to an argument of Hardy, which is only cited, not reproduced in MM's context and notation. Hardy's axioms include one (his Axiom 2) which is not a Requirement of MM, and I am not sure how it would follow from MM's Requirements. There is no discussion of this in MM. Moreover, Hardy's argument leading to the conclusion which MM cites is not clear to me.
>
> In summary, Theorem 2 seems central to MM's main conclusions. Even assuming Theorem 2, there seem other aspects of MM's argument which deserve close scrutiny and more extensive exposition than given in MM."

Dr. Müller correctly points out that the argument [result] of Hardy is a "standalone mathematical theorem" which doesn't involve Hardy's axioms. Although

measurements!

There seems a logical vagueness here which I can't quite put my finger on and makes me nervous. Nevertheless, I don't think there is any physical problem with MM's statement (on the bottom of p. 2 that "d is equal to the (affine) dimension of $\mathcal{S}$".

I've not carefully read the proof of this theorem, I have no reason to doubt it, and I agree that there would be no reason for MM to reproduce it. When I wrote the above, I was worried about the relation between Hardy's $K(N)$ and MM's d_N given the difference in their setups, and I failed to notice that the relation is irrelevant in MM's context.

References

[1] Hardy, L., Quantum Theory From Five Reasonable Axioms, arXiv:0101012v4

[2] Barrett, J., Information procession in generalized probabilistic theories, Phys. Rev. A **75** 032304 (2007), arXiv:0508211v3

[3] Masanes, L., and Müller, M., A derivation of quantum theory from physical requirements, New J. Phys. 13:063001 (2011), arXiv:1104.1483v4 , called MM above//

[4] Chiribella, G., D'Ariano, G., and Perinotti, P., Probabilistic theories with purification", Phys. Rev. A. 81, 062348 (2010), arXiv: arXiv:0908.1583, (called CDP10 below)

[5] Chiribella, G., D'Ariano, G., and Perinotti, P., Informational derivation of quantum theory, Phys. Rev. A 84,012311 (2011), arXiv: arXiv:1011.6451 (called CDP11 below)

[6] Masanes, Ll., and Müller, M., Information-theoretic postulates for quantum theory, arXiv: 1203.4516v1

[7] Zaopo, M, Informational Axioms for Quantum Theory, arXiv:1205.2306v1