

The two Bell's theorems of John Bell

Author

Wiseman, HM

Published

2014

Journal Title

Journal of Physics A: Mathematical and Theoretical

Version

Accepted Manuscript (AM)

DOI

[10.1088/1751-8113/47/42/424001](https://doi.org/10.1088/1751-8113/47/42/424001)

Rights statement

© 2014 Institute of Physics Publishing. This is the author-manuscript version of this paper. Reproduced in accordance with the copyright policy of the publisher. Please refer to the journal's website for access to the definitive, published version.

Downloaded from

<http://hdl.handle.net/10072/65174>

Griffith Research Online

<https://research-repository.griffith.edu.au>

The Two Bell's Theorems of John Bell

H. M. Wiseman

Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111, Australia

Abstract. Many of the heated arguments about the meaning of “Bell’s theorem” arise because this phrase can refer to two different theorems that John Bell proved, the first in 1964 and the second in 1976. His 1964 theorem is the incompatibility of quantum phenomena with the dual assumptions of locality and determinism. His 1976 theorem is the incompatibility of quantum phenomena with the unitary property of local causality. This is contrary to Bell’s own later assertions, that his 1964 theorem began with that single, and indivisible, assumption of local causality (even if not by that name). While there are other forms of Bell’s theorems — which I present to explain the relation between Jarrett-completeness, “fragile locality”, and EPR-completeness — I maintain that Bell’s two versions are the essential ones. Although the two Bell’s theorems are logically equivalent, their assumptions are not, and the different versions of the theorem suggest quite different conclusions, which are embraced by different communities. For realists, the notion of local causality, ruled out by Bell’s 1976 theorem, is motivated implicitly by Reichenbach’s Principle of common cause and explicitly by the Principle of relativistic causality, and it is the latter which must be forgone. Operationalists pay no heed to Reichenbach’s Principle, but wish to keep the Principle of relativistic causality, which, bolstered by an implicit “Principle of agent-causation”, implies their notion of locality. Thus for operationalists, Bell’s theorem is the 1964 one, and implies that it is determinism that must be forgone. I discuss why the two ‘camps’ are drawn to these different conclusions, and what can be done to increase mutual understanding.

1. Introduction

This special issue celebrates the 50th anniversary of both the submission and (surprisingly‡) the publication of Bell’s paper which introduced his eponymous theorem [1]. The present year is also the 50th anniversary of the submission of a less well-known paper by Bell [2], a prequel to the 1964 paper, which, due to editorial negligence, was not published until 1966 [7]. This delayed prequel already contained a powerful impossibility theorem relating to hidden variables, a theorem which was formulated independently, and in an even stronger form, by Kochen and Specker, in 1967 [8]. While the latter’s names are now irrevocably associated with this contextuality theorem, Mermin famously rescued Bell’s earlier result from obscurity in his 1993 paper “Hidden variables and the two theorems of John Bell” [7].

The present paper is also about two theorems proven by John Bell, but neither of them is that published in 1966. Rather, I will argue that, contrary to Bell’s retrospective description of his own work [5], Bell proved two different forms of (what is known as) Bell’s theorem. The first is that proven in 1964 [1], that there are quantum predictions incompatible with any theory satisfying locality and determinism (L&D). The second is that proven in 1976 [4], that there are quantum predictions incompatible with any theory satisfying local causality (LC). I use the terms *theorem* and *proven* advisedly. I am not concerned here with the intuitions that may have motivated Bell when he wrote his papers, but rather with what he rigorously proved (at least with rigour that should satisfy most theoretical physicists).

Now while LC is a strictly weaker concept than L&D (as will be discussed), it was highlighted by Fine in 1982 [9] that the range of phenomena respecting L&D is the same as the range of phenomena respecting LC. Therefore Bell’s 1976 theorem is a corollary of Bell’s 1964 theorem, as well as implying it. Since the two Bell’s theorems are thus logically equivalent, why should anyone but a pedantic historian of science distinguish them? The answer is: because the different theorems point in different directions. Bell’s 1964 theorem suggests that Bell experiments leave us with a choice: accept that physical phenomena violate determinism, or accept that they violate locality. Bell’s 1976 theorem suggests that Bell experiments leave us with no choice: we must accept that physical phenomena violate local causality.

These different interpretations of Bell’s theorem are characteristic of different communities of physicists (which could be broadly characterised as operationalists and realists respectively). The failure to distinguish *which* Bell’s theorem (1964 or 1976) one is discussing, and to use unambiguous terminology, is, I know from experience, a significant cause of miscommunication (see *e.g.* Ref. [10]). That is why I believe it is worthwhile for me to take the time to explain, and for the reader to take the time to

‡ Bell’s paper was received on 4 November 1964 and published in the November/December issue of that year [1]. Presumably this issue was not actually printed until 1965, which presumption is supported by the fact that in many [2, 3, 4, 5, 6] of his papers that referred to Ref. [1], Bell gave the year of its publication as 1965. Those with the date of 1964 may of course have been editorially “corrected”.

understand, exactly what Bell proved, when he proved it, how he defined the terms he used to describe what he had proven, and how his terminology changed over time. I also take the opportunity to examine alternate formulations of Bell’s theorems, and to explore how neither L nor LC are immediate consequences of relativity theory, but rather require at least one additional assumption, different in the two cases.

Section 2 of this paper establishes the notation and assumptions that are common to most of the paper. In Section 3 I examine the text of Bell’s 1964 paper, give my reading of it in terms of locality and determinism, but also analyse alternate readings. In Section 4 I present Bell’s 1976 paper, with its concept of local causality, and examine precedents for this concept. Section 5 looks critically at Bell’s later representations of his own 1964 and 1976 theorems. Section 6 covers related theorems by Fine and Jarrett, and introduces a new concept of “fragile locality” in order to relate Jarrett-completeness to EPR-completeness. Section 7 reviews the attitudes of the two “camps” (operationalists and realists), how these relate to Bell’s two Bell’s theorems, and to the Principles of relativity, common cause, and “agent-causation”. Finally, in Sec. 8, I discuss why we cannot avoid considering at least Bell’s two forms of Bell’s theorem, and explain my recommendations to both communities regarding how to cite, how to state, and how to construe, the appropriate Bell’s theorem, if they wish for productive discussions, or at least “peaceful coexistence” *à la* Ref. [11].

2. Notation and Assumptions

In this section I introduce the notation (largely following Bell [1, 4, 12, 13]) used throughout the paper and the assumptions which are implicit (or, in some papers, explicit) in Bell’s theorems.

2.1. Notation

Consider two spatially separated and non-communicating observers, Alice and Bob, who choose to make particular measurements, and observe their outcomes. Alice’s choice of what measurement she is to make is represented by a . This is a free choice from some set $\{a\}$. For Bell’s analysis [1] of the EPR-Bohm scenario [14], a was the direction along which to measure spin, and was notated \vec{a} . I am suppressing the vector notation for convenience and generality. The outcome of that measurement once she makes it is A . The corresponding things for Bob are indicated by b and B . For the most rigorous tests we should arrange that Alice’s choice a and outcome A are space-like separated from Bob’s b and B . Here to avoid notational complexity I am using these symbols both for the values and for the events yielding them (*i.e.* the event of Alice’s making and implementing the choice a , and the event of the measurement’s yielding the result A).

In the common past light cone of A and B is c , an agreed-upon reproducible preparation procedure for the experiment which yields A and B . I denote by λ any other variables (in addition to c , a , and b) that may affect the outcomes obtained. That

is, they are variables not specified by the preparation procedure c , and as such may be deemed ‘hidden variables’. They may be specified in any region of space-time as long as it is earlier than the choices a and b . Any space-like hypersurface may be chosen to define ‘earlier’ here. (Note that there is no need to assume, as Bell did in some papers, that λ corresponds to events in space-time. It could, for example, indicate a quantum pure state $|\psi\rangle$, which is a property of a hypersurface.)

Using the above notation, we can define a phenomenon ϕ by the relative frequencies $f_\phi(A, B|a, b, c)$, where c may be fixed, but a and b will, in the situations we are interested in, be allowed to vary. I use the term ‘relative frequencies’ rather than ‘probabilities’ to emphasise that a phenomenon is a feature of the world. However I will use the term ‘quantum phenomenon’ for a phenomenon predicted by quantum mechanics, given how empirically successful quantum mechanics is. Phenomena should be distinguished from *theories*. A theory θ for a phenomenon ϕ comprises at least the following: the set Λ of values of λ ; a mapping from c to a probability measure $d\mu_\theta(\lambda|c)$ on Λ ; and a specification of a joint probability distribution $P_\theta(A, B|a, b, c, \lambda)$, which predicts the phenomenon:

$$\int_{\Lambda} d\mu_\theta(\lambda|c)P_\theta(A, B|a, b, c, \lambda) = f_\phi(A, B|a, b, c) \quad (1)$$

Note that here, as everywhere in this paper, when an equation involving variables appears without quantifiers for those variables, a universal quantifier (\forall) is to be understood.

The set Λ could contain a single λ , or the measure $d\mu_\theta(\lambda|c)$ could be singular, or the dependences of $P_\theta(A, B|a, b, c, \lambda)$ on λ could be trivial. In these cases the theory does nothing more than reproduce the phenomenon. Such theories can be called *operational* theories. A theory, operational or otherwise, might postulate other variables or events, in addition to A , a , B , and b , which are later than the space-like hypersurface introduced to define λ . For the purpose of this paper, except for the discussion of Criterion 11, I will assume that this is not the case, and restrict consideration to those events listed.

2.2. Assumptions

Axiom 1 (Macroreality) *In an individual experimental run, exactly one outcome A and exactly one outcome B really happen, and are not ‘relative’ to anything.*

Axiom 2 (Space-time) *The concepts of space-like separation, light-cones, space-like hypersurface etc. can be applied unambiguously in ordinary laboratory situations.*

Axiom 3 (Arrow of time) *A cause can only be in the past of its effect.*

Axiom 4 (Free choice) *a and b are freely chosen (i.e. effectively uncaused), and so are independent of c , of λ , and of each other.*

While these axioms may be doubted (see Sec. 7), most discussions of Bell’s theorem take them for granted and so will I, in this paper. There may be other implicit assumptions but they would be at least as uncontroversial as the above.

Note that the term ‘cause’ in Axiom 3 is here undefined; this axiom has force only when used in conjunction with other notions explicitly involving the idea of causation. In this context, I also specify the following:

Principle 5 (relativistic causality) *In relation to causation, ‘the past’ is to be understood as ‘the past light cone’.*

Principle 6 (manifest agent-causation) *If a phenomenon exhibits an event that is statistically dependent on a freely chosen action, then that action is a cause of that event.*

I call these ‘Principles’ rather than an ‘Axioms’ because I will entertain rejecting them, or at least closely related Principles. With regard to relativistic causality, members of the ‘realist camp’ identified in Sec 1 would, indeed, typically reject this Principle in response to Bell’s theorem (see Sec. 7.3). Regarding manifest agent-causation, while Principle 6 may seem indisputable, at least one causal decision theorist [15] rejects it (R. Briggs, pers. comm.). While this Principle is not directly related to Bell’s theorem (see Sec. 7.3), there is a stronger, and therefore more disputable, notion of agent-causation (Principle 24) which the ‘operationalist camp’ seems impelled to embrace.

3. Bell’s 1964 Theorem

3.1. The “almost universal” reading

My reading of Bell’s 1964 paper [1], which is the reading, Bell later complained, “almost universally reported” [5], is that the theorem *proven* there is:

Theorem 7 (Bell 1964) *There exist quantum phenomena for which there is no theory satisfying locality and determinism.*

For theorems like this I will also use the more succinct expression: quantum phenomena violate either locality or determinism.

Bell’s meaning for determinism in a theory is, I think, uncontroversial:

Definition 8 (determinism) *A theory θ is deterministic, i.e. satisfies determinism (D), iff (if and only if) all probabilities $P_\theta(A, B|a, b, c, \lambda)$ are either zero or one.*

That is, determinism means that the outcomes are given by functions $A_\theta(a, b, c, \lambda)$ and $B_\theta(a, b, c, \lambda)$. In a more general setting, determinism means that an outcome A is determined by the “initial values” [1] of the hidden variables (since they are specified prior to some space-like hypersurface) plus any subsequent, freely chosen, measurement settings (which are uncaused events by Axiom 4).

Bell’s meaning for “locality” in a theory is, by contrast, extremely controversial. Therefore we must carefully examine how Bell defines it in this paper. In his Sec. I (Introduction), a single paragraph which also serves as the abstract, Bell states (with my edits, introduced in the interests of brevity, indicated by curly brackets and ellipses)

... Einstein, Podolsky and Rosen ... argue{ed} that quantum mechanics could not be a complete theory but should be supplemented by additional variables {in order} to restore to the theory causality and locality [2]. In this note that idea will be formulated mathematically and shown to be incompatible with the statistical predictions of quantum mechanics.

Here Bell identifies two concepts, causality and locality. He does not define causality at this (or any other) point, and the reference “[2]” in the above quote is to a footnote which consists solely of a quote from Einstein,

But on one supposition we should, in my opinion, absolutely hold fast: the real factual situation of system S_2 is independent of what is done with the system S_1 , which is spatially separated from the former.

and its source, my Ref. [16].

It is not immediately clear in the text whether Bell means this footnote to gloss “causality and locality”, or just “locality”. However, in the very next sentence Bell says

... the requirement of locality, or more precisely that the result of a measurement on one system be unaffected by operations on a distant system with which it has interacted in the past

That is, assuming that the “real factual situation” of a system is what is probed by measuring it, Bell’s definition of locality follows from the supposition of Einstein’s which he quotes. Now the notions of being “independent of what is done with” or “unaffected by operations on” a system clearly refer to the action of an agent (say Alice) on her system, and mean that Alice’s action has no statistical effect. In the context of a Bell experiment, Alice’s action on her system is to perform some particular measurement a , chosen by her. Thus we can formalise Bell’s definition of locality, provisionally (in the sense that we will have to test this definition against what Bell says elsewhere in his paper), as follows:

Definition 9 (locality) *A theory θ is local, i.e. satisfies locality (L), iff*

$$P_\theta(B|a, b, c, \lambda) = P_\theta(B|b, c, \lambda), \quad (2)$$

plus the corresponding equation $P_\theta(A|a, b, c, \lambda) = P_\theta(A|a, c, \lambda)$ for Alice.

The “corresponding equation for Alice” will go unstated in definitions from now on. Note that it is implicit in the above equation that the function $P_\theta(B|b, c, \lambda)$, which we have not previously considered, exists.

As yet, what Bell means by causality remains unexplained. However, Bell closes Sec. I with these statements:

It is the requirement of locality ... that creates the essential difficulty. ... There have been attempts ... to show that even without such a separability or locality requirement no “hidden variable” interpretation of quantum mechanics is possible. ... {However,} a hidden variable interpretation of

elementary quantum theory [5] has been explicitly constructed. That particular interpretation has indeed a grossly non-local structure. This is characteristic, according to the result to be proved here, of any such theory which reproduces exactly the quantum mechanical predictions.

This passage demonstrates a number of points with Bell's terminology. First, he uses "separability" and "locality" interchangeably (as he had done also in his earlier paper [2]), though he uses the latter more often so it is the term I have adopted. Also, rather than the two assumptions of causality and locality, he states the two assumptions of "hidden variables" and locality. In Ref. [2], Bell used "hidden variables" to mean «"dispersion free" states» which make individual measurement results "determined", and I am sure that is what he intends here also. Now Bell clearly implies that for Einstein's program, locality is the problem, not causality, because of "[5]", which is Bohm's pilot wave interpretation [17], a perfectly fine (albeit nonlocal) deterministic hidden variable theory. All of this suggests to me that by "causality" Bell means nothing more than determinism, and was, again, simply being promiscuous with his terms. That is, Bell's *theorem*, as he describes in his final sentence, is that any deterministic theory of quantum phenomena must be nonlocal.

Bell uses the terms "causality" only a single time in the paper. Elsewhere he is clear that the second assumption (*i.e.* in addition to locality) he requires for his theorem is determinism. In the last sentence of Bell's Sec. V, he states: "the statistical predictions of quantum mechanics are incompatible with separable predetermination." (In the spirit of Axiom 3, Bell makes no distinction between *predetermined* outcomes and *determined* outcomes.) The first sentence of Bell's Sec. VI (Conclusion) reiterates this statement of his theorem, but with precise descriptions rather than with any of the above terms:

In a theory in which *parameters* are added to quantum mechanics to *determine the results of individual measurements*, without changing the statistical predictions, there must be a mechanism whereby *the setting of one measuring device can influence the reading of another instrument, however remote*.

Here the italics are my addition, to highlight Bell's definitions of determinism and (the negation of) of locality.

As the above quote shows, Bell definitely means locality specifically as the absence of any influence of the *setting a* of the remote measurement device. This confirms my above reading of his definition of locality in Definition 9. In fact this reading is confirmed in two more places in the paper. In the first paragraph of Sec. II, he states it in the context of an experiment on a spin-singlet state of two spin-half particles:

we make the hypothesis [2] . . . that if the two measurements are made at places remote from one another, the orientation of one magnet does not influence the result obtained with the other.

Recall that Bell's note [2] is the quote from Einstein, which Bell reuses here, and which, as I have shown above, Bell associates with locality. Later in that same section, he again states

The vital assumption [2] is that the result B for particle 2 does not depend on the setting a of the magnet for particle 1, nor A on b .

Finally, in (my) Ref. [2], Bell’s prequel to his 1964 theorem, his use of locality is consistent with the above, as he elucidates the “grossly nonlocal character” of Bohm’s theory [17] by showing that “in this theory . . . the disposition of one piece of apparatus affects the results obtained with a distant piece.”

With this understanding of Bell’s assumptions, the structure of his paper is as follows. His Section II is titled “Formulation”, and this is where he gives the *mathematical formulation* (as promised in his Introduction) of the dual assumptions of causality (i.e. determinism) and locality, his equation (1),

$$A(a, \lambda) = \pm 1, B(b, \lambda) = \pm 1.$$

In the context (spin measurements), the existence of these functions and their possible values, does indeed capture these dual assumptions§.

Now prior to this formulation, Bell gives a one-paragraph motivation for considering hidden variable theories, by reiterating the EPR argument in its spin-singlet version [18]. The crucial sentence, which follows Bell’s statement of the assumption “the orientation of one magnet does not influence the result obtained with the other” (quoted above), is

Since we can predict in advance the result of measuring any chosen component of $\vec{\sigma}_2$ by previously measuring the same component of $\vec{\sigma}_1$, it follows that the result of any such measurement must actually be predetermined.

Here Bell has made a mistake. His conclusion (predetermined results) does not follow from his premises (predictability, and the hypothesis stated in the preceding sentence). This is simple to see from the following counter-example. Orthodox quantum mechanics (OQM) is a theory in which the setting a of one device does not statistically influence the result B obtained with the other:

$$P_\theta(B|a, b, c, \lambda) = P_\theta(B|b, c, \lambda). \quad (3)$$

Here, if c were to correspond to preparation of a mixed quantum state ρ_c , the variable λ would allow for a pure-state decomposition; in purely operational QM, only c would appear||. But in OQM it is of course not true that the results of spin measurements are predetermined for a singlet state as Bell is considering.

Now, as stated, Bell’s first paragraph (not counting the abstract) serves only to motivate the formulation of the theorem. This is to be expected from its placement prior to the mathematical formulation. It is also reflected in how the paragraph ends: “this predetermination implies the possibility of a more complete specification of the state.” Having thus introduced the *possibility* of hidden variable (i.e. deterministic)

§ In this paper, Bell does not use c , which is why it is absent as an argument in these functions. However λ can always be assumed to include c for the purpose of characterising a theory as local or deterministic.

|| In the present case, where ρ_c is a singlet state (which is pure), there is no distinction.

theories, he then formulates such theories, in the next paragraph, under the hypothesis of locality, as per his equation (1):

Let this more complete specification be effected by means of parameters λ
 The result A of measuring $\vec{\sigma} \cdot a$ is then determined by a and λ , and {similarly} the result B . . . , and {his equation (1)}. The vital assumption [2] is that the result B for particle 2 does not depend on the setting a of the magnet for particle 1, nor A on b .

The end of this quotation is the locality assumption already quoted above. Finally, that Bell’s EPR paragraph forms no part of his 1964 theorem (which in this paper he calls his “result”) is clear from the fact that after he has *formulated* the problem in Sec. II, and *illustrated* it in Sec. III, he finally gets to Sec. IV where “The main result will now be proved.”

Thus I would classify Bell’s mistake in this paragraph as a peccadillo, having no impact on the main result in his paper. It would have been an easy mistake for Bell to have made, if he had the idea that EPR had already proven determinism from some sort of locality assumption, and did not think hard about whether it was the same as the locality assumption he was about to use in his own theorem. Indeed the paper could be made completely sound simply by replacing “it follows” in the above (“Since we can predict . . .”) quote by “the obvious explanation is”, or “EPR’s premises imply”. Although Bell believed that he was reproducing EPR’s argument, EPR’s premises (which are never stated by Bell) are *not* equivalent to locality (as defined here by Bell), and they *do* justify the conclusion of predetermined outcomes; see Ref. [19] and Appendix A.2.

3.2. Alternate Readings

As the beginning of the preceding subsection implies, in later life Bell represented his 1964 paper differently from how it was “almost universally reported”, as I will discuss in Sec. 5. There is also a community of Bell’s followers [20, 21, 22] who support his stance; see Sec. 7.2. Their reading of Bell’s paper differs from the usual understanding as follows. First, they see the first paragraph of Bell’s “Formulation” section as an essential part of his 1964 theorem, the first part of a two-part argument. Second, they see this first part as legitimately deriving determinism from predictability and the assumption of locality, since they see Bell’s concept of locality as being more general than my reading of it. Basically, they see Bell’s notion of locality as being local causality (see next section), which requires an outcome to be statistically independent not only of the remote setting, but also of the remote outcome. In particular, they do not see operational quantum mechanics as a counter-example whereby predictability and locality do not imply determinism. Thus they represent Bell’s 1964 theorem as being that quantum phenomena violate the assumption of locality.

This reading of course has the advantage that it does not require the reader to assume that Bell made a mistake, however inconsequential, in his epitome of the

EPR argument. It also has another advantage (Maudlin, pers. comm.): Bell’s second sentence, which talks of Einstein’s hope that hidden variables would “restore to the theory causality and locality [2]” also implies that Bell believed that OQM did not satisfy the locality condition, or at least that Bell believed that Einstein believed this. If this is meant to reflect Bell’s beliefs then it is, on my reading, the same mistake he made in reproducing the EPR argument, and is equally unimportant here. Regarding Einstein’s beliefs, I will return to these in Sec. 3.3, since Bell’s definition of locality does not exactly coincide with the quote from Einstein he uses here.

To me, the advantages of this reading are demonstrably outweighed by its many disadvantages: i) it does not explain why Bell would, in 1964, state his result *four times* as requiring two assumptions, locality and determinism, and not once as requiring only the assumption of locality; ii) it does not explain why in his first subsequent paper on the topic of hidden variables [12], after seven years to think about how best to explain his result, he still states it (somewhat redundantly) as being “that no local deterministic hidden-variable theory can reproduce all the experimental predictions of quantum mechanics” [12]; iii) it does not explain why Bell would, in 1964, define locality *four times* in terms of independence from the remote setting, as per Definition 9, and never any other way¶; iv) it does not explain why Bell would state the conclusion of the supposedly crucial first part of his theorem as being merely that it “implies the possibility of a more complete specification of the state.”; v) it does not explain why Bell would place this supposedly crucial first part *prior* to the mathematical formulation of his result, and not mention it anywhere else in the paper.

In short, to claim that the theorem Bell proved in his 1964 paper rules out a unitary ‘localistic’ hypothesis (to coin a word), despite the fact that he never articulates a hypothesis that would be adequate to the task, and invariably states his result as also requiring the assumption of determinism, is to do violence to the notions of theorem and proof. One might be tempted to venture that Bell never intended to provide a rigorous argument from precisely stated premises to the sought-for contradiction with quantum mechanics, and so any one is free to take his heuristic argument and make it rigorous in whatever way seems best. This is unfair to Bell, as he certainly does give such a rigorous argument, from the premises of D and L. Furthermore, Bell clearly did intend his argument to be rigorous: having previously <found wanting {others’} *attempts* to show that ... no “hidden variable” interpretation of quantum mechanics is possible>, he contrasts these attempts with “the result to be *proved* here” (my emphasis).

Of course one is free to construct a proof using the EPR paragraph of Bell’s paper to derive L and D, if one corrects it by beginning with a sufficiently strong localistic hypothesis H. This may be useful to obtain a pedagogical proof of Bell’s theorem (in a general sense), as in Ref. [23] for example, but that should not be confused with Bell’s

¶ In particular, there is no evidence to support the suggestion (Norsen, pers. comm.) that Bell began with a general notion of locality, along the lines of local causality, and only narrowed to this definition after he had established determinism via his EPR paragraph. Bell’s statement of the locality hypothesis in the middle of that paragraph is exactly the same as elsewhere in the paper.

1964 theorem, or its proof. In this context, one could ask what assumption H Bell may, hypothetically, have had in mind, which prompted him to give the motivating EPR argument, even if he failed to articulate it, for whatever reason. Readers not interested in this exploration may skip to Sec. 4.

3.3. Hypothetical hypotheses

In this subsection, I try to identify a localistic hypothesis H which is unambiguous, is sufficiently powerful to make the EPR argument work, is not tantamount to assuming D and L, and plausibly could have been held by Bell in 1964. As we will see, even with this strictly limited aim, there are difficulties (in particular due to the final clause) with all three candidates that present themselves.

The first obvious candidate for H is local causality (LC), a notion which Bell later adopted, as will be discussed in the following section. Although the notion of LC had been formulated mathematically by the logician Boole as early as 1854 (see Ref. [24]), and by the philosopher Reichenbach in 1956 [25], there is no reason to suspect that it was known to Bell in 1964⁺. In physics, the concept was developed incrementally, by Bell and others, in the first half of the 1970s (see Sec. 4.1). The two papers which Bell references in his EPR paragraph, Einstein’s autobiographical notes [16], and the paper on the “Paradox of Einstein, Rosen, and Podolsky” (*sic.*) by Bohm and Aharanov [18] both use the same sort of localistic notion Bell does, referring only to the influence of the remote setting. The quote from Ref. [16] has already been examined* and will be further discussed below. The assumptions of Ref. [18] are wholly implicit, but the only notion of locality there is in the negative: a hypothetical nonlocal “hidden interaction” which “would have to be instantaneous, because the orientation of the measuring apparatus {for A } could very quickly be changed, and the spin of B would have to respond immediately to the change.”

The second candidate for H is *completeness*, as EPR [29] defined it, or, rather, characterised it, using a web of several conditions, necessary or sufficient as required; see Ref. [19] and Appendix A.3 of the current paper. As shown there, the notion of completeness is sufficient to derive a contradiction with the predictions of quantum mechanics, using the arguments in Bell’s 1964 paper. The EPR paper was

⁺ Intriguingly [26], Reichenbach — whose “principle of common cause” for correlations [25] is, given the axioms of Sec. 2.2, equivalent to LC [27] — had previously considered (and dismissed) hidden variable theories in 1944. That he did not combine these areas of knowledge to prove Bell’s 1976 theorem indeed “goes to show how subtle and deep was Bell’s insight” [26].

* It must be admitted, however, that if Bell had chosen a *different* quote from Ref. [16], *viz.* “the real situation of S_2 must be independent of what *happens to* S_1 ” — my emphasis — then my case would not have been so clear cut; see Sec. 7.2 for the nuances of word choice in this context. One should not read too much into Einstein’s diction in the quote in this footnote, however. Writing about the same time on “quantum mechanics and reality” [28], Einstein elevates the “principle of contiguity”, saying that without it empirical physics would be impossible. Einstein’s definition of this principle is practically the same as the other quote which Bell used to motivate his definition of locality: “{For} objects far apart in space (A and B): external influence on A has no direct influence on B.”

unquestionably known to Bell, and he even entitled his paper “On the Einstein Podolsky Rosen Paradox”. However, as just indicated, the EPR notion of completeness is far from simple in logical terms, and there is no evidence that Bell had a formal understanding of it. He makes no mention of any of the formal criteria EPR introduce, and does not use “complete” in their technical sense, talking, for example, of a “more complete specification of a state.” He never quotes from the EPR paper [29], preferring Einstein’s less technical notes [16], and in his prequel [2], the latter was his *only* reference for the “Einstein-Podolsky-Rosen paradox.” Bell’s other inspiration in 1964, the paper by Bohm and Aharonov [18], is even less formal than Ref. [16]. Despite its title, it is disconcertingly disengaged from the EPR paper. It does not avail itself of EPR’s terminology, mentioning neither “completeness” nor “reality”, and ubiquitously gives the three authors’ names in the wrong order.

The final candidate for H is the supposition Bell quotes from Ref. [16], already mentioned above. It is worth quoting it here:

Supposition 10 (No telepathy) *The real factual situation of system S_2 is independent of what is done with the system S_1 , which is spatially separated from the former.*

(For the name I have given this supposition, see the quote below.) Now although Bell seemed to indicate (twice) that this was equivalent to his definition of locality, it is different in that it requires not that Bob’s *result* B be independent of Alice’s setting a , but rather that the “real factual situation” of Bob’s system be thus independent $\#$. Einstein never states what qualifies something to be a “real factual situation”. With some thought, one can propose a strong but reasonable formalisation for this concept such that the assumption (which, prior to the passage quoted above, Einstein makes explicitly) that systems have real factual situations, plus the no telepathy supposition, has the same force as local causality (see Criterion 11). However, there is no suggestion of such a formalisation in Ref. [16]. Nor does Einstein use it there to make the argument that Bell wants to make, from predictability to determinism. Even if Bell had worked out such a formulation from studying Einstein’s text, Supposition 10 alone would not be strong enough to do its hypothetical role in Bell’s EPR paragraph. As Bell would have known if he had made such a study, Einstein readily admitted that the prior assumption mentioned above is necessary:

One can escape from this conclusion [that statistical quantum theory is incomplete] only by either assuming that the measurement of S_1 (telepathically) changes the real situation of S_2 or by denying independent real situations as such to things which are spatially separated from each other. Both alternatives appear to me equally unacceptable.

Leaving aside any hypothetical formulation of of Einstein’s “real factual situation” that could have worked for Bell, it is interesting to ask the question whether

$\#$ In this, Supposition 10 arguably has a similar relation to locality as per Definition 9 as does local causality to factorizability; see Sec. 4.1. For a couple of reasons, which this footnote is too small to contain, I suggest “separability” as an appropriate name for the exact analogue to local causality.

Supposition 10 alone is violated by OQM (as Bell implies). Einstein certainly argues in Ref. [16] that this is so, if, as “it appears to me”, “one may speak of the real factual situation of the partial system S_2 ” even when one takes OQM to be complete. That is, in Einstein’s understanding, if quantum mechanics is complete then the quantum state of S_2 , conditioned on measurements on S_1 , must be a “real factual situation”. A ‘realist’ understanding of OQM does involve such a physical collapse of the wavefunction, which requires specifying some foliation of space-time and does indeed violate Einstein’s supposition. However strictly operational QM does not violate Supposition 10, because strict operationalists do indeed operate “by denying independent real situations as such”, as Einstein subsequently allowed. Thus even if Bell did have “no telepathy” in mind for locality, his implication that OQM violates locality would be incorrect, because OQM can be interpreted purely operationally. It is also important to remember that OQM satisfies locality as Bell actually expresses it, as per Definition 9, regardless of whether one has a realist or an operationalist understanding of the theory.

4. Bell’s 1976 theorem

As mentioned above, in 1971 Bell wrote another (*i.e.* in addition to Ref. [2]) review of the “hidden-variable question” [12], this time with the opportunity to give his own result its due emphasis. With characteristic modesty, he co-referenced his own 1964 paper with Refs. [30, 31], which closely followed his own presentation, relying upon (deterministic) “hidden variables” with the “locality” condition or assumption. As will be discussed below, the presentation in Ref. [12] does show some evolution in Bell’s thinking, but a clear conceptual advance is not to be found until his 1976 paper [4], where he introduced the term “local causality” (LC).

Reference [4] made use of the concept “local beable”, Bell’s term (introduced in Ref. [32]) for elements of reality that were assumed to be localisable to points, or at least small regions, in space-time. Many potential elements of reality are not usually conceived this way — a quantum state, for example, does not exist in space at all, but rather is defined on a space-like hypersurface [21]. As Bell himself said in 1981 [5], “These variables $\{(\lambda)\}$ could well include, for example, quantum mechanical state vectors, which have no particular localization in ordinary space time.” However, one could, if one wished, imagine that every point on that space-like hypersurface carries a representation of the quantum state, and in that way one could construct an ontology of the sort Bell wished for in 1976. Thus we can continue to use Bell’s earlier notation of Ref. [1], to which he returned in his definitive presentation on the subject [13], just before he died. The conceptual advance which the notion of local beables prompted for Bell was to treat settings and outcomes (and other, postulated, beables) on the same footing.

Thankfully, unlike the case for Ref. [1], it is not necessary to do a detailed textual analysis to convince anyone of the premise Bell introduced in 1976††. He first mentions

†† That said, one may become confused if one tries to follow his argument from his equations and figures

the idea that theories of classical physics exhibit *local determinism*. Being formulated in terms of beables, this is more general than his considerations in 1964, but it certainly implies the joint premise of L&D he made there. Then he says

We would like to form some notion of *local causality*, in theories which are not deterministic, in which the correlations prescribed by the theory, for the beables, are weaker.

Using the usual notation, Bell’s new concept can be characterised as follows:

Criterion 11 (local causality) *A theory θ is locally causal, i.e. satisfies local causality (LC), only if*

$$P_{\theta}(B|A, a, b, c, \lambda) = P_{\theta}(B|b, c, \lambda). \quad (4)$$

Note that this is a necessary criterion (‘only if’, not ‘iff’) because here Bell’s general notion is applied only to the specific beables relating to hidden variables (λ) prior to some particular space-like hypersurface, and settings and outcomes posterior to it.

The idea of local causality is that any events, such as A and a , which lie outside the past light cone of event B , are statistically irrelevant to B . Of course, one would only expect LC to hold if one’s theory included a description of all prior variables λ relevant to the outcomes; otherwise outcome A could contain information about λ which would make it statistically relevant to outcome B . Note, however, that the presence of λ does not imply any assumption of determinism here. That is, the concept of LC may apply even for non-deterministic theories such as OQM. Indeed, when ρ_c is an unentangled state, there is always an instance of OQM (corresponding to a decomposition of ρ_c into pure product states [33]) which satisfies LC. But for entangled states, orthodox *quantum mechanics is not locally causal*, as Bell says as his 3rd section heading of Ref. [4]. Moreover, using assumption (4), one can (and Bell does) derive an inequality which contradicts quantum predictions, thus proving

Theorem 12 (Bell 1976) *There exist quantum phenomena for which there is no theory satisfying local causality.*

To put it another way, “quantum mechanics is *not* embeddable in a locally causal theory” [4].

It is unfortunate (for me writing this article, and for anyone reading Ref. [4] subsequent to his more famous 1964 paper) that, almost as soon as he had introduced the term LC, Bell started using the term “locality” for the same thing, referring to a “locality inequality” (for what everyone else calls a Bell inequality) and the “locality hypothesis” in Ref. [4]. However, at least in his last word on the subject [13], Bell showed his preference unequivocally. Here Bell carefully discusses various answers to the question “what cannot go faster than light?”, and when referring to the concept 11, he uses the term “local causality” at least 17 times, and the term “locality” only twice

alone, since he introduces in his Eq. (2) the notion of local causality as the statistical irrelevance of beable B , but applies it to his Eq. (5) as the statistical irrelevance of beables M and B .

(in a single paragraph discussing experiments which is rather disconnected from the rest of the paper)†. Thus I believe it respects Bell’s final will (and avoids confusion when reading his pre-1976 papers) to use only the term LC for this concept, and to ignore Bell’s occasional lapses in this regard.

4.1. Precedents for Bell’s concept of Local Causality

While Bell certainly introduced, and promoted, the term LC for the concept 11, there is a brief history, in the context of quantum phenomenology, of similarly formulated mathematical conditions. In particular, in 1974 Clauser and Horne introduced the notion of a “factored form” [34] of the joint probability distribution:

Definition 13 (factorisability) *A theory θ is factorisable, i.e. satisfies factorisability (F), iff*

$$P_{\theta}(A, B|a, b, c, \lambda) = P_{\theta}(A|a, c, \lambda)P_{\theta}(B|b, c, \lambda). \quad (5)$$

Equation (5) is mathematically equivalent to Eq. (4) (with the corresponding equation for Alice of course). Clauser and Horne describe it as “a reasonable locality condition” and an “extrapolation of the common-sense view that there is no action at $\{a\}$ distance.” They go on “{W}e call any theory in which it holds an objective local theory.¹⁵” A principal result of their paper is that “the predictions of objective local theories and of quantum mechanics differ.”

On the basis of the above, one could argue that Theorem 12, which I have called the 1976 Bell theorem, should be called the 1974 Clauser-Horne theorem. However, Bell considered his formulation to be more fundamental, saying [13]

Very often . . . factorization is taken as the starting point for the analysis. Here we have preferred to see it not as the *formulation* of “local causality”, but as a consequence thereof.

In the sense that Bell defined local causality generally, in terms of beables rather than outcomes and settings *etc.*, Bell’s point is an important one. Bell’s terminology is also to be preferred, as “locally causal” has a unique associated noun (local causality), whereas “objective local”, comprising two adjectives, does not, and instead suggests two separate assumptions, objectivity and locality. The conceptual advance of Bell’s 1976 theorem lies in the unitarity of the notion of local causality, so I believe it is appropriate to give the credit to Bell.

As a final point, Clauser and Horne themselves say, in their footnote (15) to the above quote, «The class of OLT . . . is essentially the class of stochastic hidden-variable theories “with a certain local character” considered by Bell (Ref. 4).» Here “Ref. 4” refers to Bell’s 1971 paper [12], so the credit for the theorem of Clauser and Horne

† He does however use, quite deliberately, the term “locally explicable” as synonym for “locally causal” with reference to correlations. It would be a mistake to regard the “explicable” here as inessential. The concept of LC is all about explaining correlations, as Bell goes to great lengths to explain in Ref. [13] and elsewhere. The idea of “explaining” correlations is formalised in Principle 25.

arguably lies with Bell in any case. Indeed, Bell’s 1971 paper gives the general expression for a correlation function allowing (in a footnote) for “indeterminism with a certain local character”. Using the notation of Sec. 2.1, Bell’s expression is

$$\langle AB \rangle_{a,b,c} = \int_{\Lambda} d\mu_{\theta}(\lambda|c) \bar{A}_{\theta}(a, \lambda, c) \bar{B}_{\theta}(b, \lambda, c), \quad (6)$$

where \bar{A}_{θ} and \bar{B}_{θ} some real-valued functions specified by the theory, which give the *average* over local indeterminism not described by λ . Since a or b can correspond to arbitrary observables, including the identity, the above characterisation does capture all possible statistics of the measurement results A and B in a theory “with a certain local character”. This phrase is, unfortunately, not fit to appear in a theorem, and Bell offers nothing better in Ref. [12]. Thus, again, it seems better to stick with Bell’s far more elegant notion of LC, and to keep my nomenclature for Theorem 12.

5. Later Commentary by Bell

As noted in Sec. 3.2, with the benefit of hindsight, Bell represented the logic of his 1964 paper [1] rather differently from how it appears to me. He had, as just mentioned in Sec. 4.1, considered “indeterminism with a certain local character” in 1971, and in a commentary from 1972 [35] he wrote:

It can ... be shown that the quantum mechanical correlations cannot be reproduced by a hidden variable theory even if one allows a “local” sort of indeterminism. For example, one could imagine that the indeterminism might be introduced at every space-time point and allowing the result of each throw to influence physical events in the future light cone of the point in question. This would not work: the quantum mechanical correlations are too perfect to permit any such statistical slop. Popper remarks {in a volume published in 1971} that he does not find this point manifest in my own paper on the subject, but it is there — very briefly.

Here Bell’s “own paper” is Ref. [1], and the implication is that Bell believed he covered the «“local” sort of indeterminism» in that paper. Presumably the “very brief” manifestation Bell has in mind is his argument from locality as per Definition 9 to determinism via predictability — the incorrect argument I discussed in Sec. 3.

In the concluding paragraph of his 1976 paper, Bell writes:

This paper has been an attempt to be rather explicit and general about the notion of locality, along lines hinted at in previous publications {[1, 12, 35, 32]}.

With regard to the later references, this quote is more than fair, as the “certain local character” of Ref. [12] did capture the implications of the notion of LC, as discussed in Sec. 4.1. With regard to Bell’s 1964 paper [1], it is also true that the explicit notion of locality (L) used there was not as *general* as LC. But the latter was *hinted at* only in so far as: (i) it is related to (though much more general than; see Appendix A) the notions

used by EPR, to whom Bell attributes his pseudoderivation of D from predictability; (ii) *unlike* Bell’s 1964 explicit notion (L), it *does* allow one to derive D from predictability. Point (ii) is easy to show: If $P(A|a, B, b, c) = 1$ or 0, then the same must be true for all $P_\theta(A|a, B, b, c, \lambda)$.[‡] But if we assume LC (Criterion 11) it follows that $P_\theta(A|a, c, \lambda) = 1$ or 0 also, which is L&D.

In any case, it seems that once Bell had explicitly defined LC, he wished all previous localistic notions he had used, in particular the notion of locality as per Definition 9, to be forgotten. Moreover, after a few years he became convinced that it was the notion of LC that he had in mind all along. In 1981 Bell says (Ref. [5], text and footnote 10 stitched together)

It is important to note that to the limited degree to which *determinism* plays a role in the EPR argument, it is not assumed but *inferred*. What is held sacred is the principle of “local causality” — or “no action at a distance”. . . . There is a widespread and erroneous conviction that for Einstein (and his followers) determinism was always the sacred principle. (My own first paper on this subject {[1]} starts with a summary of the EPR argument from locality to deterministic hidden variables. But the commentators have almost universally reported that it begins with deterministic hidden variables.)

Thus, Bell implies in 1981 that both he and Einstein were always using the notion of LC, which Bell characterises later in this 1981 paper in the same way (Criterion 11) as in 1976. As argued in Sec. 3, there is only one plausible reading of “locality” in Bell’s 1964 paper, and it is not LC. Nor did EPR formulate such a general notion, Refs. [20, 22] notwithstanding.

Let this be absolutely clear: I am not accusing Bell or his followers of intellectual dishonesty. Local causality is such an elegant notion that in hindsight it is astonishing that it had to wait until 1976 (or at least 1974 [34]) to be applied to quantum phenomena. If EPR had stated this assumption in 1935, they could have used their example to formulate a far simpler argument showing OQM to be an unsatisfactory theory (by their lights). Indeed, Bell credits them with exactly this argument, in Sec. 6.8 of Ref. [13]§. Furthermore LC seems to me a more natural localistic notion than L, as will be discussed in Sec. 8. The fact that Bell thought, mistakenly, that one could obtain determinism from locality (as defined by him in 1964) suggests strongly that he already had then an intuitive localistic notion (*e.g.* “no action at a distance” [5]) that was different from, and stronger than, L. From personal experience I know that it is easy to forget, after a decade or two, the details of a published argument, and to remember the intuitive idea which in fact only crystallised in one’s mind later. I have no doubt that anyone familiar with Bell’s later work could have deduced from Bell in 1964 the precise notion of LC in

[‡] Strictly, this need not hold for those λ such that $d\mu_\theta(\lambda|c) = 0$, that is, for those λ which never occur in the experiment. We are not concerned with such λ in definitions of D, L, or LC.

[§] Of course EPR would not necessarily have endorsed the name local causality in this context because their aim was to prove the incompleteness of OQM. That is why they built localistic notions into their definition of completeness; see Appendix A.2.

Criterion 11, with little effort on either’s part. But hypotheticals do not alter history, and the theorem Bell *proved* in his 1964 paper is not the same as Theorem 12 from 1976.

6. Related Theorems

The interpretation of Bell’s theorem has generated a large literature. In this section I briefly review some of the results that will be relevant to my Discussion section.

6.1. Fine

As discussed above, Clauser and Horne in 1974 [34] recognised that factorisability (5) — which is a consequence of LC — is no more general than Bell’s 1971 “certain local character” [12]. And in that paper Bell noted that the same formula (6) applied to correlations under this constraint as under the constraint of L&D. That is,

Theorem 14 (Fine 1982) *A phenomenon respects LC iff it respects L&D.*

Note that this theorem is not saying that LC and L&D are logically equivalent concepts. Recall that there are theories that satisfy LC but that do not satisfy L&D, such as OQM for product states.

I call the above Fine’s theorem, even though his paper [9] came much later than the above, because it was he who focussed attention on it, rather than burying it in footnotes, as in Refs. [34, 12]. Fine refers to “factorizable stochastic models” (*i.e.* F of Definition 13) rather than to LC, but in terms of the observable events they are equivalent. Fine duly credits Clauser and Horne [34] with Theorem 14, but also establishes it directly, for the case of two settings per party, each with two outcomes [9]. The ‘if’ part of Theorem 14 is trivial, since L&D implies F. But F does not imply L&D, so the non-triviality is in the ‘only if’ part. For a general proof, see Appendix A of Ref. [36]. The basic idea, as in Ref. [12], is that in a factorizable theory, any stochasticity in an outcome can be *reproduced* by having it determined by additional hidden variables that may be assumed to reside locally.

6.2. Jarrett / Shimony

Fine’s theorem showed that, in terms of its experimental consequences, LC might as well be thought of as L&D. Not long after, Jarrett [37] showed that LC is in fact logically equivalent to a subtly different conjunction. He (and, independently, Shimony [38]) introduced a new concept,

Definition 15 (Jarrett-completeness) *A theory satisfies Jarrett-completeness (JC) iff*

$$P_{\theta}(B|A, a, b, c, \lambda) = P_{\theta}(B|a, b, c, \lambda). \quad (7)$$

Using this he proved

Theorem 16 (Jarrett 1984) *Factorizability (or “strong locality” [37]) is equivalent to the conjunction of:*

- (i) *Locality*||, also known as “parameter independence” [38].
- (ii) *Jarrett-completeness*, also known as “outcome independence” [38].

The corollary, via Theorem 12, is

Corollary 17 (Bell 1976 as L&JC) *Quantum phenomena violate either locality, or Jarrett-completeness.*

Here my theorem-naming convention indicates that Bell’s 1976 argument would have gone through with the assumptions of L&JC rather than the assumption he did make there (LC); the same convention applies to Theorems 22, 23, and 35 below.

Orthodox quantum mechanics of course satisfies L but violates JC, while deterministic hidden-variable theories (such as Bohm’s theory [17]) satisfy JC but violate L. Thus the options presented by Theorem 17 (Jarrett’s version of Bell 1976) are resolved in the same way as in Theorem 7 (Bell 1964) for these two leading approaches to QM (of those that satisfy the axioms of Sec. 2.2). This is not surprising because JC can be thought of as a weak version of determinism:

Lemma 18 *Determinism implies Jarrett-completeness.*

Shimony [38] called the violation of L “controllable nonlocality”, which is perhaps fair enough. However it is another matter to state that such violations “provide (at least in principle) the means for superluminal signal transmission,” as does Jarrett [37] (and, in much the same words, also Shimony [40]). For example, although Bohm’s theory violates locality, its statistical predictions are identical to OQM, and so it does not allow superluminal signalling. This is because the value of a particle’s position (a hidden variable in Bohm’s theory) cannot be known to a macroscopic agent any better than in OQM. In some approaches to Bohm’s theory [41] — although not others [42] — this lack of knowledge is inevitable. It is only for *operational* theories that locality is the same as signal-locality [43],

Definition 19 (signal-locality) *A phenomenon respects signal-locality iff $f(B|a, b, c) = f(B|b, c)$.*

See Ref. [44] for further discussion of the distinction between L and signal-locality.

It is important to note, again *contra* Jarrett [37], that only signal-locality is an absolute “requirement of relativity theory”, which can be seen as follows. Let us assume, in addition to Axioms 1–4, Principle 6, for manifest agent-causation. Now since Definition 19 involves relative frequencies, any violation of signal-locality would

|| Jarrett [37] agrees with Definition 9. So does Howard [39]. But the latter uses the term “separability” for Jarrett-completeness, a usage which for me is not intuitive, and which has not caught on. Recall that Bell used the terms “separable” and “local” interchangeably in his earliest papers. One might be tempted to call Jarrett-completeness “causality” because of Theorem 16 (since factorisability amounts to local causality here) but it is clear that these concepts are already overloaded with proposed names.

mean, by Principle 6, that a is a cause for B . But because a is not in the past light-cone for B , Principle 5 (relativistic causality), would be violated. By contrast, if a *theory* violates L, but not signal-locality, these assumptions do not entail that a causes B , because the *phenomenon* exhibits no dependence of B on a . Hence there is no conflict with relativistic causality. I will revisit this point in Sec. 7.3.

Jarrett naturally calls JC merely “completeness”; I have antepended his name because this concept is neither implied by, nor implies, EPR-completeness, $\text{Com}(\theta)$, as I have analysed it (see Appendix A.2). In Jarrett’s defence, however, one can relate them with only a weak additional assumption as follows. First I will define

Definition 20 (fragile locality) *A theory satisfies fragile locality (FL) iff*

$$P_\theta(B|a, b, c, \lambda) \in \{0, 1\} \implies P_\theta(B|a, b, c, \lambda) = P_\theta(B|b, c, \lambda). \quad (8)$$

That is, if, for some values of its arguments, there exists a function $B_\theta(a, b, c, \lambda)$, then that function does not depend upon a . The terminology “fragile” is appropriate because it depends upon exact determination, which is a set of measure zero in correlation-space. Note that Jarrett’s “strong locality” (*i.e.* factorizability, or, more or less, LC) implies locality, which in turn implies fragile locality.

Given Definition 20, it is not difficult to show (see Appendix A.4) that

Theorem 21 *If EPR-completeness implies X , then Jarrett-completeness plus fragile locality implies X .*

(Indeed, it could be argued that fragile locality is one of the localistic notions that EPR implicitly assume; see Appendix A.2.) A corollary of this, via Theorem 35, is

Corollary 22 (Bell 1964 as FL&JC) *Quantum phenomena violate either fragile locality, or Jarrett-completeness.*

Note that the joint assumption of fragile locality and Jarrett-completeness does *not* imply LC. Thus, this version (Corollary 22) of Bell’s 1964 theorem, unlike Bell’s original version, does not have stronger premises than Bell’s 1976 theorem. That is, it is not subsumed by Theorem 12 (using LC) in the way that Theorem 7 (using L&D) is. The same remarks also apply to Theorem 35.

Like Theorem 35, the statement of Bell’s 1964 theorem as Corollary 22 can only be proven for quantum phenomena with perfect correlations (as Bell considered in 1964). Thus it does not say anything about phenomena that are experimentally accessible; see Appendix A.3 for details. This shortcoming in the theorem can be removed by strengthening FL to L (giving Theorem 17), or JC to D, giving

Corollary 23 (Bell 1964 as FL&D) *Quantum phenomena violate either fragile locality, or determinism.*

I call this a version of Bell’s 1964 theorem because $\text{D\&FL} \iff \text{D\&L}$ (even though only $\text{L} \implies \text{FL}$). In fact one might deem Bell’s characterisation of locality in Ref. [1] to be too imprecise to distinguish between L of Definition 9 and FL of Definition 20, but to avoid further terminological complication I will not pursue that line of argument.

7. The Two Camps

As stated in the Introduction, those who ponder the implications of Bell’s theorem, (physicists, philosophers, and, increasingly, information scientists), and who accept Axioms 1–4, can broadly be grouped into two camps: operationalists and realists. Note that these four Axioms respectively rule out: relativist approaches such as Everett’s relative state interpretation [45] (popularly known as the many worlds interpretation [46]); giving up free choice (albeit maybe to only a modest degree [47]), which Bell called “superdeterminism” [13]; ubiquitous wormholes, as discussed in Ref. [48] (pp. 481–3); and retrocausal ideas [49, 50].

7.1. Operationalists

The operationalist attitude is particularly common among quantum information theorists, and is well represented by this quote from a recent text-book [51]:

{To derive Bell’s theorem, we} now make two hypotheses about the behaviour of the composite system:

Hidden variables. We assume that the results of any measurement on an individual system are predetermined. Any probabilities we may use to describe the system merely reflect our ignorance of these hidden definite values, which may vary from one experimental run to another.

Locality. We assume that Alice’s choice of measurement does not affect the outcomes of Bob’s measurements, and *vice versa*.

As we have seen, these two hypotheses more or less capture the point of view advocated by EPR.

Note that Schumacher and Westmoreland use exactly the same terminology (*hidden variables*, *predetermined*, and *locality*) as Bell’s 1964 paper, and I think their presentation very fairly captures the definitions in that paper. As for its implication, they state

From a philosophical point of view, Bell’s theorem and the experimental confirmation of quantum theory constitute one of the most remarkable results in all of physics. The dual hypotheses of hidden variables and locality have a pretty good claim to the “common sense” view of the world. Nevertheless, *at least one of them must be wrong*.

Another famous quantum information text-book with a similar presentation is Ref. [33]. Showing some appreciation of the literature, they say

Vast tomes have been written analyzing ... the subtly different assumptions which must be made to reach Bell-like inequalities. Here we merely summarise the main points.

But their summary is very similar to that of Ref. [51]. In place of “hidden variables” or “predeterminism” they state

The assumption that ... physical properties ... have definite values ... which exist independently of observation. This is sometimes known as the assumption of *realism*.

Having then stated the locality assumption (practically identically to Ref. [51]), they go on:

These two assumptions together are known as the assumption of *local realism*. They are certainly intuitively plausible assumptions about how the world works, and they fit our everyday experience. Yet the Bell inequalities show that at least one of these assumptions must be incorrect.

It is worth making a few remarks about this terminology. “Local realism” is certainly a common term these days for what Bell experiments disprove. As far as I can tell, it was introduced by d’Espagnat in 1979 [52], although it is reminiscent of the term “objective local” used by Clauser and Horne [34]. However, the term “realism” was never used by Bell, and is generally used synonymously with (pre)determinism, as Mermin (another early proponent) admits [53]. It is because of this that I have called Axiom 1 — the independent existence of detectors and results — “macroreality”. (For further discussion of the infelicity of “local realism”, see Ref. [54].) By my casual observation, “realism” is used more often by those who see it as the obvious hypothesis to reject, such as Nielsen and Chuang [33]:

Most physicists take the point of view that it is realism that must be dropped from our worldview in quantum mechanics.

Because operational QM satisfies locality, it is natural for operationalists to advocate espousing locality, and rejecting the second assumption required for Bell’s 1964 theorem. Nevertheless I find it a surprising development in the language surrounding Bell’s theorem, that this second assumption should often be called “realism”. The usual philosophical meaning of “realism” is the belief that entities exist independent of the mind, a worldview one might expect to be foundational for scientists. However, there is an influential sub-community of operationalists, the most vocal among which go by the name of QBists (Quantum Bayesianists), who reject precisely that, saying [55] “reality differs from one agent to another.” For QBists, reality is identified with the experience of a single agent, and they dismiss Bell’s theorem by emphatically rejecting Axiom 1 (macroreality). Each QBist believes in his own reality, but refuses to

assign correlations, spooky or otherwise, to space-like separated events, since they cannot be experienced by any single agent. Quantum mechanics is thus explicitly local in the QBist interpretation.

As Louis XIV might have said, «L’état quantique, c’est moi.»

7.2. Realists

Since operationalism, taken to its logical extreme in QBism, is anti-realist in all senses of the word, it seems appropriate to use the term “realists” for those at the other end

of the spectrum in the debate on the meaning of Bell's theorems. This camp, in which we can locate Bell himself, even if he was never so dogmatic as some [20, 21, 22], is characterised by two convictions: (1) that correlations need to be explained (Bell [5] calls this nothing less than “the scientific attitude”); (2) that nature should have a unified description, in which anthropocentric notions such as “detector settings” should play no fundamental role.

Despite the name I have given it, the realist camp does not advocate that we should cleave to “reality” [33, 53] (that is, determinism, also expressed as “hidden variables” [51, 7]) and spurn the locality of Definition 9. Rather, its members generally read Bell's theorem as requiring (in addition to the axioms of Sec. 2.2) only a single assumption [20, 21, 22], and often explicitly reject the two-assumption formulation of the operationalist camp. A quote from the textbook of Maudlin [21] (p. 19) will suffice:

Bell himself derived the result as part of an examination of so-called local hidden-variable's theories {which add} parameters whose values determine the results of experiments. Bell's results are therefore sometimes portrayed as a proof that local deterministic hidden variables theories are not possible.

This is a misleading claim. It suggests that the violation of the inequality may be recovered if one just gives up determinism or hidden variables. But as we have seen, the only assumption needed to derive the inequality is that the result of observing one particle is unaffected by the experiment carried out on the other ... a condition {which} is generally called “locality”.

At first sight this seems in blatant conflict with the operationalist camp who claim that an additional assumption to locality, which I would call determinism, *is* necessary to derive Bell inequalities. But on closer inspection, Maudlin's definition of locality is subtly different. It refers not to the effect of “Alice's choice of measurement” [51], but to Alice's “experiment” (or, a few pages earlier, Alice's “observation”) which presumably includes the outcome she gets as well as her choice of measurement. That is, by “locality”, Maudlin means “local causality” as introduced by Bell in 1976.

The preference for the 1976 theorem was Bell's own, as already noted, and the Conclusion to his final paper on the matter expresses the realist view well [13]:

The obvious definition of “local causality” does not work in quantum mechanics, and this cannot be attributed to the “incompleteness” of that theory. ...

Do we then have to fall back on “no signalling faster than light” as the expression of the fundamental causal structure of contemporary theoretical physics? That is hard for me to accept. For one thing we have lost the idea that correlations can be explained ... More importantly, the “no signalling {faster than light}” notion rests on concepts which are desperately vague, or vaguely applicable. The assertion that “we cannot signal faster than light” immediately provokes the question

Who do we think *we* are?

Bell clearly expresses the two tenets of the realist camp, as listed in the first paragraph of this section, but here gives greater weight to the second. While “no signalling faster than light” is not the same as locality (see Sec. 6.2), the objection to using anthropocentric concepts (*i.e.* actions by agents) as the basis for physical theory applies equally to both.

The important point is that the notion of LC does not rely on action by agents: a theory could violate LC with no agent choice at all. For example, consider a single quantum particle in a pure state split between two boxes, one sent to Alice and the other to Bob. Then simultaneously (by some foliation) they open the boxes, and observe whether the particle is there or not. Clearly their observations will be perfectly correlated (the particle is seen by one and only one of them) but in OQM there is no variable λ that factorizes the joint probability distribution as per Eq. (5); there is no variable that determines which box the particle will be found in. Thus OQM can be shown to violate LC without measurement choice. (This is sometimes known as the “Einstein Boxes” argument [56].) The subtlety of Bell’s insight is that the only *phenomena* that violate LC *do* require agency (free choice of settings).

7.3. Irreconcilable differences?

If one has ambitions to reconcile these two camps, then first one must understand the reasons for the differences between their presentations.

The first, and perhaps most important, difference is in terminology, as I have already discussed. As I have shown, exemplars in Bell’s writings exist for both, competing, definitions of “locality”. The operationalist camp use it as Bell used it 1964 and a few subsequent papers, and the realist camp as he used it, occasionally, from 1976 onwards, as a synonym for LC. The second difference is that the two camps then, as expected, state Bell’s theorem as Theorems 7 and 12 respectively.

Being followers of Bell, the realist camp must be aware that, from 1976 onwards, Bell used the term “local causality” in preference to “locality”, especially in formal expositions. While some prominent realists sometimes promote Bell’s preferred term [23], it is more common to see “locality” used for LC. For example, the above quote from Maudlin’s 1994 book [21] is unchanged in the 2011 edition, while I must admit that in an earlier work [57] I mentioned “local causality” only to say that it was “often abbreviated to locality,” which I thereafter used. Presumably this substitution is made because “locality” is (a) shorter and thus more memorable, (b) a single word and thus not able to be confused with a dual concept; and (c) simpler and thus suggestive of a natural concept. My present stance is that these advantages do not outweigh the disadvantages: (1) it disregards Bell’s final and most thorough communication on the subject [13]; and (2) it makes it appear as if the two camps are in disagreement over basic logic, when in fact that is not the case.

The realist camp might object along these lines: «Why should we change our terminology? Let the operationalist camp use the phrases “parameter independence” [38] or “setting independence” if they need a term for such an unnatural notion.» There

are two answers to that. First, the operationalists (being, by nature, primarily interested in the applications of quantum mechanics) are generally less well read in this area of metaphysics, and are less likely to be familiar with such alternate phrases. Second, the one piece of literature everybody reads (or at least references) with regards Bell's theorem is his 1964 paper, and there the operationalist (or indeed anyone without the "benefit" of hindsight) will read Definition 9 for "locality" in Bell's words no fewer than four times, as well as a similar concept in Einstein's words.

For the operationalist, locality is a natural assumption because OQM respects it, and because it follows from the assumptions of Sec. 2.2, *if* one strengthens Principle 6 to the (still reasonable)

Principle 24 (agent-causation) *If, in a theory, an event is statistically dependent on a freely chosen action, then, in that theory, that action is a cause of that event.*

Holding to this principle would be, as far as I can tell, the simplest way for the operationalist camp to justify L from relativity (Principle 5), and thereby conclude, from Theorem 7 (Bell 1964), that D must be violated[¶]. While realists would not necessarily reject Principle 24, they would not see it as an appropriate one for fundamental analysis of the physical world, because of the central role it gives to "freely chosen action" (see the extended quote from Bell in the preceding section).

The challenge for the the operationalist camp is to recognise that there is another localist notion (I speak of LC of course) which OQM does *not* respect, and which is not only meaningful, but which also follows from the assumptions of Sec. 2.2 (with or without Principle 6), *if* one holds to another reasonable principle:

Principle 25 (common cause [25]) *If two events with no direct causal relation are correlated, then they have a common cause κ , such that conditioning on κ will eliminate the correlation. This is so even if the correlation is already conditional, as long as the conditioning event is not an effect of both events.⁺*

Just as local causality \implies locality \implies signal locality, we have for these principles

Theorem 26 *common cause \implies agent-causation \implies manifest agent-causation.*

The less obvious (the first) of these implications can be seen as follows. If an event A is statistically dependent upon a free choice a , that means that they are correlated. Now a is uncaused by Axiom 4, so by Principle 25 the only option is that a causes A , as in Principle 24.

[¶] Alternatively, since signal-locality is a matter of fact about quantum phenomena (and the world, as far as we know), it would presumably be possible to derive L using some sort of "no fine-tuning" Principle (*i.e.* that the causal structure of the theory should not change under small changes in the probability distribution for the hidden variables). There is a similar argument deriving LC from signal-locality, Reichenbach's Principle 25, and "no fine-tuning" [58].

⁺ With regard to the last clause, Axiom 3 implies that as long as an event (say f) is in the past light cone of event e , it cannot be an effect of e . Moreover, if Principle 5 is also assumed, then unless f is within the future light cone of e , it cannot be an effect of e .

For the realist camp, Principle 25 is fundamental (correlations need to be explained), and is used implicitly to motivate LC (rather than L) from relativity. Thus, since Theorem 12 (Bell 1976) rules out LC, members of the realist camp would typically reject Principle 5. That is, they would allow space-like causal influences, as long as they respect Axiom 3 according to some foliation of space-time. As an aside, it seems to me that, for this purpose, one could use, in place of Principle 25, the rather simpler

Principle 27 (causal completeness) *The probability of an event, conditioned on its causes, is unchanged if further conditioned on other events that are not its effects.*

To close, the above challenge to operationalists evokes a third reason the realist camp should not try to own “locality”: (3) by using “local causality” and referencing Bell’s 1976 paper, realists might induce operationalists to read something else by Bell, and to appreciate the power of Criterion 11 and Theorem 12. That said, operationalists would still not be compelled to accept space-like causal influences, as they could follow Ref. [27] and reject Principle 25 (common cause) rather than Principle 5 (relativistic causality). This is simply a sophisticated way of saying that operationalists could still, in Bell’s words [59], “shrug off a correlation.”

8. Discussion

Bell’s theorem is the most profound ramification of quantum theory that has been experimentally confirmed. In this paper I have presented several different formulations of it: Theorems 7, 12, 17, 22, 23, and 35. However, I would argue that there are only two essential forms with which quantum physicists should be familiar, corresponding to the two Bell’s theorems of my title, Theorem 7 (Bell 1964), and Theorem 12 (Bell 1976). Even though either theorem implies the other, the two have independent roles in quantum foundations, as the expressions of Bell’s insight preferred by operationalists and realists respectively.

Of these I side with the realists in thinking that Bell’s 1976 theorem is the more important. The assumption of local causality is strictly weaker than the assumption of locality and determinism, so Bell’s 1976 theorem subsumes Bell’s 1964 theorem. Also, the notion of local causality seems to me the most natural expression of the spirit of special relativity for statistical theories: that, given a complete (within the theory) specification of the physical situation prior to Alice’s and Bob’s choices of measurement settings, Alice’s outcomes are statistically independent of space-like separated events (such as Bob’s choice and his outcomes). Then, as Bell proved, “quantum mechanics is not locally causal and cannot be embedded in a locally causal theory” [4]. That is, orthodox quantum mechanics violates local causality, and adding hidden variables cannot alleviate the problem.

I have argued strongly (in Secs. 4 and 7.3) for using the term “local causality”, not “locality”, in this statement of Bell’s theorem for two principal reasons. First, it follows the formal definition Bell proposed in 1976 [4], and the usage he resolutely

promoted in his ultimate presentation [13]. Second, and more importantly, it is necessary to avoid the confusion (and even hostility) which can arise when people read Bell’s 1964 paper — or other seminal papers such as Refs. [30, 60, 61, 37] — and see there “locality” defined in a weaker (*i.e.* less general) sense, in which Alice’s outcomes are required only to be statistically independent of Bob’s choice of measurement settings. Another reason to endorse the terminology “local *causality*” is that, under Axioms 1–4 and relativistic causality (Principle 5), local causality is simply the expression of Reichenbach’s “Principle of common cause” [25], Principle 25.

If (as I am advocating) Bell’s promulgation of the term “local causality” should be respected, then why not also his abandonment of his earlier notion of locality? Unfortunately for Bell, one cannot choose which contribution one will be remembered for. Bell’s 1964 paper is, and probably always will be, the most famous of his papers, the one to which new researchers in the field will turn as a primary source. It is thus hard to argue that “locality” as per Definition 9 should not be taken seriously as a concept when it played such a foundational role. Also, locality can be derived from assumptions 1–5, together with the reasonable Principle 24 which I have tentatively called “agent-causation”.

Bell’s 1964 theorem — that quantum phenomena are incompatible with either locality or determinism — is naturally favoured by operationalists because operational quantum mechanics *does* respect locality, leaving determinism as the obvious assumption to forgo. Rather than determinism (or “hidden variables” or “realism”), some have suggested a weaker assumption (see Sec. 6.2) which can be combined with locality in order to derive a contradiction with the predictions of quantum mechanics (Theorem 17). I do not believe such formulations shed much light beyond Bell’s original formulation, any more than I believe Theorems 23 (replacing locality with a weaker assumption) or 22 (replacing both locality and determinism with weaker assumptions) do.

For operationalists, Bell’s 1964 theorem is not only more natural, it is also more useful, in the following sense. As an operationalist, one is interested only in theories in which the hidden variables (λ) are, in principle, knowable to some agent. Such theories must be local if signal locality is to hold even when the knowledge of all agents is taken into account. Similarly, such theories must be deterministic if the measurement outcomes are to be predictable when taking into account the knowledge of all agents. Thus, if one believes in signal locality as a matter of principle, then Bell’s 1964 theorem implies that quantum phenomena must be unpredictable in principle [62]. (See Ref. [63] for a more formal presentation.) In information science, guaranteed unpredictability or randomness is a valued resource for numerical algorithms and security protocols [64]. It is thus of interest that a Bell experiment can be turned into a protocol that generates (or, rather, expands) randomness, certified by the principle of signal-locality [64]. Moreover, because Alice’s and Bob’s results are *correlated* in Bell-type experiments, it is possible, in principle, to use them to obtain *secret shared randomness* between distant parties, which is even more valuable, enabling communication with security guaranteed by signal

locality [65]*.

While distributing a secret key is the most important proposed application of Bell’s inequalities, there are more specialised applications of quantum correlations that are, in themselves, a proof of violation of local causality, without statements about secrecy being required. In particular, there are protocols — such as those enabling teams of accused robbers [67] or crooked accountants [68] to lie to their interrogators without fear of discovery — in which *all* correlations are perfect‡, and the appearance of predetermination is so strong that it is hard not to feel that they must be nonlocal. Nevertheless, according to the terminology I recommend, neither of Bell’s theorems allow us to conclude that quantum phenomena are nonlocal, taking that word to mean the negation of “local”.

In conclusion, for a proper appreciation of the foundational importance of Bell’s theorem to physics, information science, and the philosophy of causation, one should be familiar with both the 1964 Bell’s theorem and the 1976 Bell’s theorem, even though they are logically equivalent. The former proves that quantum phenomena are either nonlocal (in a “causation by agents” sense) or undetermined, while the latter proves that quantum phenomena violate local causality (in a “common cause for correlations” sense). For those who prefer the latter theorem, as Bell ultimately did, and who find the verb phrase “violate local causality” lacking in pith, I make one final terminological suggestion, regarding a pre-existing term:

Definition 28 *‘Bell-local’ is an acceptable synonym for ‘locally causal’, and ‘Bell-nonlocal’ for its negation.*

This usage has a long history (going back at least to 1981 [70], in the negative) and is still very much current (see *e.g.* the review [71]). It is not a term which Bell would ever have used, and, because of Bell’s inconsistent use of “locality”, it risks confusing unwary readers of the ‘historical’ literature on Bell-nonlocality††. But, on the positive side, it is surely a term of which the latter-day Bell would have approved.

Acknowledgements

I am deeply grateful for comments on drafts of this manuscript by, in chronological order of first communication: Eric Cavalcanti, Eleanor Rieffel, Michael Hall, Matthew

* In fact, other assumptions are also necessary to guarantee security, in particular that neither Alice’s nor Bob’s laboratory inadvertently leaks any kind of information. The proof in Ref. [65] requires perfect correlations (as do the protocols in the subsequent paragraph), but the idea can be applied in quantum key distribution even with imperfect correlations, as follows. If one assumes the correctness of quantum mechanics (a strictly stronger assumption than signal locality), then violating a Bell inequality may allow the distribution of a secret key without Alice or Bob having to trust the entanglement source or their detectors [66]. Similar results apply for randomness expansion [64].

‡ The class of quantum phenomena which violate local causality in this way has been called “quantum pseudo-telepathy” [69], an echo of Einstein’s term (see Supposition 10 and the subsequent discussion).
 †† For example, “Bell’s locality assumption” in Ref. [61] does not mean Bell-locality as per Definition 28, but rather (as it should, since Aspect *et al.* cite only Bell’s 1964 paper [1]), locality as per Definition 9.

Pusey, Rachael Briggs, Roger Colbeck, Travis Norsen (who wishes it to be noted that he completely disagrees with my analysis and finds my arguments flimsy, circular, and wholly unconvincing), Tim Maudlin, and Mitchell Porter. This research was supported by the ARC Centre of Excellence CE110001027 and by the ARC Discovery Project DP140100648.

Appendix A. EPR, Completeness, Bell’s Theorem, and fragile locality

In this Appendix, I summarise the formalisation in Ref. [19] of the criteria used by EPR [29]. The purpose of this material is three-fold. The first is to explain how, at least on my reading of EPR, Bell’s 1964 theorem can be thought of as a proof that no theory of quantum phenomena can be EPR-complete. Thus, EPR’s completeness is a candidate for Bell’s imagined hypothesis H in his 1964 paper. The second is so that the reader will appreciate how subtle EPR-completeness is. This is to explain, in part, why it is implausible to imagine that Bell had this concept in mind. The third is to relate EPR-completeness to Jarrett-completeness, which forms the last part of the Appendix. This Appendix draws some of its material directly from Ref. [19].

Appendix A.1. Preliminaries

To connect the formalism of Bell’s papers to the language of EPR, it is necessary to talk of *systems*. To give this an operational meaning (that is, one grounded in macroscopic events), assume that events a and A are locatable within a space time region α which is disjoint from another region β containing b and B . We can then identify Alice’s and Bob’s systems with α and β respectively. While EPR, surprisingly, never mention spatial separation of their systems (let alone space-like separations) they do assume that their systems are “no longer interacting”. For ease of comparison with Bell’s work I will take this to hold because they are space-like separated.

EPR also refer to *physical quantities* pertaining to systems α and β . I will denote these by $\hat{a} \in \mathbb{P}_\alpha$ and $\hat{b} \in \mathbb{P}_\beta$. I use a hat because in quantum mechanics physical quantities are represented by operators, but there is no implication here that quantum mechanics must be correct. Because there are many ways to measure a physical quantity, \hat{a} is associated with an equivalence class of settings: $a \in \mathbb{S}_{\hat{a}}$, and similarly for \hat{b} and b . In what follows, unless otherwise specified, the symbol $\forall a$ is to be understood as meaning all possible a ; that is, $\forall a \in \mathbb{S}_{\hat{a}}$ for some $\hat{a} \in \mathbb{P}_\alpha$; and likewise for b , A , and B , *mutatis mutandis*. I will assume that c is fixed throughout.

Appendix A.2. Completeness

As with most of the concepts they introduce, EPR explain completeness reasonably precisely:

{T}he following requirement for a complete theory seems to be a necessary one:
every element of the physical reality must have a counterpart in the theory.

Here, and — except where noted — below, the italics are as in the original. From a careful examination of EPR’s text [19], one finds that their criterion for completeness, for a theory θ , should be expressed using formal logic as

Criterion 29 (EPR-completeness)

$$\text{Com}(\theta) \implies \left[\forall \hat{b} \in \mathbb{P}_\beta, \mathcal{EPR}(\hat{b}|c) \implies \mathcal{Rep}_\theta(\hat{b}|c) \right].$$

Here I have used $\mathcal{Rep}_\theta(\hat{b}|c)$, standing for ‘is represented’, to denote EPR’s concept that \hat{b} “has a counterpart” in theory θ , and $\mathcal{EPR}(\hat{b}|c)$ to mean that the property \hat{b} is an ‘element of physical reality’.

EPR do not actually define what it means for \hat{b} to ‘have a counterpart in the theory’, but I think it is uncontroversial to take it to imply that the outcome of any measurement of \hat{b} is determined in the theory:

Criterion 30 (Representation in the theory)

$$\mathcal{Rep}_\theta(\hat{b}|c) \implies \forall B, b, \lambda, P_\theta(B|b, \lambda, c) \in \{0, 1\}.$$

In other words, $\mathcal{Rep}_\theta(\hat{b}|c)$ implies that there exists a function $B_\theta(b, \lambda, c)$. Note that saying this is strictly stronger than saying that a measurement of \hat{b} has predetermined outcomes, which would allow for a functional dependence on a , the measurement setting for the other system: $B_\theta(b, a, \lambda, c)$. For an explanation as to why EPR cannot have intended a to appear as an argument in their notion of representation in the theory, see Ref. [19]. However, this explanation does not actually rule out the alternative criterion where the consequent in Criterion 30 is replaced by $\forall a, B, b, \lambda, P_\theta(B|a, b, \lambda, c) \in \{0, 1\}$. Interestingly, adding the assumption of fragile locality, as per Definition 20 in Sec. 6.2, would immediately transform this alternative criterion into Criterion 30.

Regarding elements of physical reality, EPR give the following sufficient criterion:

If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.

To enable a non-disturbing measurement on β they consider an *indirect* measurement (i.e., a measurement on the other system α), so formally we have:

Criterion 31 (Element of Physical Reality)

$$\mathcal{EPR}(\hat{b}|c) \iff \exists a : [\neg \mathcal{Dist}(\beta|a, c) \wedge \mathcal{Pre}(\hat{b}|a, c)].$$

Here $\mathcal{Dist}(\beta|a, c)$ means that the measurement a disturbs the system β , while $\mathcal{Pre}(\hat{b}|a, c)$ means that the measurement a makes the value of \hat{b} predictable.

Like representation, predictability can be defined uncontroversially:

Definition 32 (Predictability)

$$\mathcal{Pre}(\hat{b}|a, c) \iff \forall B, b, A, f(B|A, a, b, c) \in \{0, 1\}.$$

Note the use here of the operationally determined relative frequencies f , not the theory probabilities P_θ — this is what differentiates predictability from determinism; see *e.g.* Ref. [63]. The second new notion — disturbance — is the most problematic in EPR’s paper, and it was the notion of disturbance which for Bohr [72, 73] was the key to refuting EPR, as he rejected their notion and proposed his own, clearly stated, sufficient condition [19]. Regardless of how EPR understood disturbance intuitively, it is apparent from the form of their argument that they thought it obvious that a measurement on system α would not disturb system β if the systems were no longer interacting. Since throughout we are taking α and β to be space-like separated, no disturbance is (for EPR but not for Bohr) a property of all indirect measurements:

Assumption 33 (No Disturbance)

$$\forall a, \neg \text{Dist}^{\text{EPR}}(\beta|a, c)$$

Appendix A.3. From Completeness to Bell’s Theorem

If we accept the criteria given by EPR as set out above, and EPR’s assumption 33, we can follow EPR’s argument to conclude that

$$\text{Com}(\theta) \wedge [\exists a : \mathcal{P}\text{re}(\hat{b}|a, c)] \implies \mathcal{R}\text{ep}(\hat{b}|c). \quad (\text{A.1})$$

As noted above, $\mathcal{R}\text{ep}(\hat{b}|c)$ is equivalent to the local predetermination of all outcomes B , which is what Bell claimed to have derived in 1964 from the assumption of locality. Here we see that $\text{Com}(\theta)$ works, where locality does not. Clearly EPR-completeness involves localistic notions, in the functional dependencies of $\mathcal{R}\text{ep}(\hat{b}|c)$ and $\mathcal{E}\mathcal{P}\mathcal{R}(\hat{b}|c)$ and in the assumption that $\neg \text{Dist}(\beta|a, c)$, but EPR never explicitly appeal to locality as defined by Bell in 1964, or to local causality as defined by Bell in 1976. At the end of the paper, EPR entertain, but then reject, the possibility of allowing $\mathcal{E}\mathcal{P}\mathcal{R}(\hat{b}|a, c)$:

{T}his point of view ... makes the reality of P and Q {(properties \hat{b} of the first system)} depend upon the process of measurement {(a)} carried out on the first system, which does not disturb the second system in any way. No reasonable definition of reality could be expected to permit this.

While this quote certainly suggests that EPR would consider locality as per Definition 9 to be a *sine qua non*, it is not an explicit assumption which they use in conjunction with other assumptions to derive their result. Rather, it is built into the very structure of their criteria [19].

If we say that, in assuming EPR’s notion of completeness, we accept all of EPR’s criteria and assumptions, then we can, following their argument, prove their theorem:

Theorem 34 (EPR) *There exist quantum phenomena for which OQM is not an EPR-complete theory.*

Moreover, assuming completeness in this way, for the singlet state and projective measurements we can follow Bell’s argument to derive that any spin observable, for

either party, must be an \mathcal{EPR} , and hence that it must have a counterpart in the theory, that is, be locally predetermined. But Bell's theorem says that any theory satisfying L&D fails to predict certain correlations on the singlet state. Thus we could state

Theorem 35 (Bell 1964 as EPR-completeness) *There exist quantum phenomena for which there is no EPR-complete theory.*

Note that this statement of Bell's theorem can only be proven for a joint system with perfect correlations, as in a spin-singlet state with projective measurements. Without perfect predictability, EPR's network of conditions literally lead nowhere. But perfect correlation is an idealisation that is never seen in experiment. Hence, while Theorem 35 is formally valid, one cannot conclude from it that there exist *phenomena in the world* for which there is no EPR-complete theory. The same criticism applies to Bell's original proof of theorem 7 in 1964. However the assumptions Bell made for this theorem, L&D rather than completeness, enables other proofs that are not subject to this criticism. That is, the assumption of L&D does give testable predictions even when the correlation is not perfect, as proven later by Clauser *et al.* [30].

Appendix A.4. EPR-Completeness versus Jarrett-Completeness

In this section I prove Theorem 21, that whatever is implied by assuming $\mathcal{Com}(\theta)$ is also implied by assuming that θ respects Jarrett-completeness and fragile locality.

First recall that, as in Eq. (A.1), there is only one implication of assuming $\mathcal{Com}(\theta)$, namely

$$\forall \hat{b} \in \mathbb{P}_\beta, [\exists a : \mathcal{Pre}(\hat{b}|a, c)] \implies \mathcal{Rep}_\theta(\hat{b}|c) \quad (\text{A.2})$$

Using Criteria 32 and 30, this means

$$\begin{aligned} \forall \hat{b} \in \mathbb{P}_\beta, [\exists a : \forall A, B, b, f(B|A, a, b, c) \in \{0, 1\}] \\ \implies [\forall B, b, \lambda, P_\theta(B|b, \lambda, c) \in \{0, 1\}]. \end{aligned} \quad (\text{A.3})$$

I now show the same implication follows from JC&FL, being more careful with quantifiers than in the main text.

Conjoining Jarrett-completeness,

$$\forall A, a, B, b, \lambda, P_\theta(B|A, a, b, c, \lambda) = P_\theta(B|a, b, c, \lambda),$$

and fragile locality

$$\forall a, B, b, \lambda, [P_\theta(B|a, b, c, \lambda) \in \{0, 1\} \implies P_\theta(B|a, b, c, \lambda) = P_\theta(B|b, c, \lambda)],$$

we have

$$\forall A, a, B, b, \lambda, [P_\theta(B|A, a, b, c, \lambda) \in \{0, 1\} \implies P_\theta(B|b, c, \lambda) \in \{0, 1\}].$$

Now it is true by definition that

$$\forall A, a, B, b, \lambda, [f(B|A, a, b, c) \in \{0, 1\} \implies P_\theta(B|A, a, b, c, \lambda) \in \{0, 1\}],$$

so we have

$$\forall A, a, B, b, \lambda, [f(B|A, a, b, c) \in \{0, 1\} \implies P_\theta(B|b, c, \lambda) \in \{0, 1\}]. \quad (\text{A.4})$$

We can weaken Eq. (A.4) to

$$\forall \hat{b} \in \mathbb{P}_\beta, \forall a, [[\forall A, B, b, f(B|A, a, b, c) \in \{0, 1\}] \implies [\forall B, b, P_\theta(B|b, c, \lambda) \in \{0, 1\}]],$$

which is the same as Eq. (A.3). In other words, the only implication that from EPR's network of criteria, Eq. (A.3), can be derived from the assumption that θ satisfies JC and FL. This completes the proof.

References

- [1] Bell J S 1964 *Physics* **1** 195 (Reproduced in Ref. [74].)
- [2] Bell J S 1966 *Rev. Mod. Phys.* **38** 447–452 (Reproduced in Ref. [74].)
- [3] Bell J S and Nauenberg M 1966 The moral aspects of quantum mechanics *Preludes in Theoretical Physics* ed de Shalit A, Feshbach H and van Hove L (Amsterdam: North Holland) p 279 (Reproduced in Ref. [74].)
- [4] Bell J S 1976 *Epistemological Lett.* **9** 11–24 (Reproduced in Ref. [74].)
- [5] Bell J S 1981 *Journal de Physique Colloques* **42** 41–62 (Reproduced in Ref. [74].)
- [6] Bell J S 1986 EPR correlations and EPW distributions *New Techniques and Ideas in Quantum Measurement Theory (Ann. New York Acad. Sci vol 480)* ed Greenberger D M (New York) p 263 (Reproduced in Ref. [74].)
- [7] Mermin N D 1993 *Rev. Mod. Phys.* **65** 803–815
- [8] Kochen S and Specker E P 1967 *J. Math. & Mech.* **17** 59–87
- [9] Fine A 1982 *Phys. Rev. Lett.* **48** 291–295
- [10] Werner R F and respondents Guest post on Bohmian mechanics URL <http://tjoresearchnotes.wordpress.com/2013/05/13/guest-post-on-bohmian-mechanics-by-reinhard-f-werner/>
- [11] Shimony A 1978 *International Philosophical Quarterly* **18** 3–17
- [12] Bell J S 1971 Introduction to the hidden-variable question *Foundations of Quantum Mechanics* (New York: Academic) pp 171–81 (Reproduced in Ref. [74].)
- [13] Bell J S 1990 La nouvelle cuisine *Between Science and Technology* ed Sarmeljijn A and Kroes P (Elsevier/North-Holland) pp 97–115 (Reproduced in Ref. [74].)
- [14] Bohm D 1951 *Quantum Theory* (New York: Prentice-Hall)
- [15] Joyce J M 2007 *Synthese* **156** 537–562
- [16] Einstein A 1949 Autobiographical notes *Albert Einstein: Philosopher-Scientist* vol 7 ed Schilpp P A translation (the odd-numbered pages) by the editor.
- [17] Bohm D 1952 *Phys. Rev.* **85** 166–193
- [18] Bohm D and Aharonov Y 1957 *Phys. Rev.* **108** 1070–1076
- [19] Wiseman H M 2013 *Annals of Physics* **338** 361–374
- [20] Dürr D, Goldstein S and Zanghì N 1992 *J. Stat. Phys.* **67** 843–907
- [21] Maudlin T 1994 *Quantum Non-Locality and Relativity* (Oxford: Blackwell)
- [22] Norsen T 2006 *AIP Conference Proceedings* **844** 281–293
- [23] Norsen T 2011 *American Journal of Physics* **79** 1261–75
- [24] Pitowsky I 1989 From George Boole to John Bell – the origins of Bell's inequalities *Bells Theorem, Quantum Theory and Conceptions of the Universe* ed Kafatos M (Kluwer Acad. Publ.) pp 37–49
- [25] Reichenbach H 1956 *The Direction of Time* (Berkeley: University of Los Angeles Press.)
- [26] Cavalcanti E G and Lal R On modifications of Reichenbach's principle of common cause in light of Bell's theorem arXiv:1311.6852v1 [quant-ph]
- [27] van Fraassen B C 1982 *Synthese* **52** 2538

- [28] Einstein A 1948 *Dialectica* **2** 320–342 translated in Ref. [75], pp. 168–173
- [29] Einstein A, Podolsky B and Rosen N 1935 *Phys. Rev.* **47** 777–780
- [30] Clauser J F, Horne M A, Shimony A and Holt R A 1969 *Phys. Rev. Lett.* **23** 880–884
- [31] Wigner E P 1970 *American Journal of Physics* **38** 1005–1009
- [32] Bell J S 1973 Subject and object *The Physicist's Conception of Nature* ed Mehra J (Dordrecht: Riedel) pp 687–690 (Reproduced in Ref. [74].)
- [33] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge, England: Cambridge University Press)
- [34] Clauser J F and Horne M A 1974 *Phys. Rev. D* **10** 526–35
- [35] Bell J S 1972 *Science* **177** 880–881
- [36] Hall M J W 2011 *Phys. Rev. A* **84** 022102
- [37] Jarrett J 1984 *Noûs* 569–89
- [38] Shimony A 1984 Controllable and uncontrollable non-locality *Foundations of Quantum Mechanics in the Light of New Technology* ed *et al* K (Physical Society of Japan) pp 225–230
- [39] Howard D 1985 *Studies in History and Philosophy of Science* **16** 171–201
- [40] Shimony A 2000 New aspects of Bell's theorem *Quantum Reflections* ed Ellis J and Amati D (Cambridge University Press) pp 136–164
- [41] Wiseman H M 2007 *New Journal of Physics* **9** 165–176
- [42] Valentini A 2002 *Physics Letters A* **297** 273–278
- [43] Skyrms B 1982 *Philosophy of Science* **49** 43–50
- [44] Dickson W M 1998 *Quantum Chance and Nonlocality* (Cambridge: Cambridge University Press)
- [45] Everett H 1957 *Rev. Mod. Phys.* **29** 454–462
- [46] DeWitt B and Graham R N (eds) 1973 *The Many-Worlds Interpretation of Quantum Mechanics* Princeton Series in Physics (Princeton: Princeton University Press)
- [47] Hall M J 2010 *Physical Review Letters* **105** 250404
- [48] Holland P 1995 *Quantum Theory of Motion: Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics* (Cambridge University Press)
- [49] Pegg D T 1980 *Physics Letters A* **78** 233–234
- [50] Price H 2008 *Studies In History and Philosophy of Modern Physics* **39** 752–761
- [51] Schumacher B and Westmoreland M 2010 *Quantum processes, systems, & information* (Cambridge, England: Cambridge University Press)
- [52] d'Espagnat B 1979 *Scientific American* **241** 158–181
- [53] Mermin N D 1980 *Phys. Rev. D* **22** 356–361
- [54] Norsen T 2007 *Foundations of Physics* **37** 311–340
- [55] Fuchs C A, Mermin N D and Schack R 2013 An introduction to QBism with an application to the locality of quantum mechanics arXiv:1311.5253
- [56] Norsen T 2005 *American Journal of Physics* **73** 164
- [57] Wiseman H M 2006 *Contemp. Physics* **47** 79–88
- [58] Wood C J and Spekkens R W 2012 The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning arXiv:1208.4119v1
- [59] Bell J S 1980 *Comments on Atom. Mol. Phys.* **9** 121–126 (Reproduced in Ref. [74].)
- [60] Aspect A, Grangier P and Roger G 1981 *Phys. Rev. Lett.* **47** 460–463
- [61] Aspect A, Dalibard J and Roger G 1982 *Phys. Rev. Lett.* **49** 1804–1807
- [62] Masanes L, Acín A and Gisin N 2006 *Phys. Rev. A* **73** 012112
- [63] Cavalcanti E G and Wiseman H M 2012 *Foundations of Physics* **42** 1329–1338
- [64] Pironio S, Acín A, Massar S, Boyer de la Giroday A, Matsukevich D N, Maunz P, Olmschenk S, Hayes D, Luo L, Manning T A and Monroe C 2010 *Nature* **464** 1021–1024
- [65] Barrett J, Hardy L and Kent A 2005 *Phys. Rev. Lett.* **95** 010503
- [66] Acín A, Brunner N, Gisin N, Massar S, Pironio S and Scarani V 2007 *Phys. Rev. Lett.* **98** 230501
- [67] Jacobs K and Wiseman H M 2005 *Am. J. Phys.* **73** 932

- [68] Sudbery A 2007 *Am. J. Phys.* **75** 720
- [69] Brassard G, Broadbent A and Tapp A 2005 *Foundations of Physics* **35** 1877–1907
- [70] d’Espagnat B 1981 *Foundations of Physics* **11** 205–234
- [71] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2013 Bell nonlocality arXiv:1303.2849
- [72] Bohr N 1935 *Nature* **136** 65
- [73] Bohr N 1935 *Phys. Rev.* **48** 696–702
- [74] Bell M, Gottfried K and Veltman M (eds) 2001 *John S. Bell on the Foundations of Quantum Mechanics* (Singapore: World Scientific)
- [75] Einstein A, Born H and Born M 1971 *The Born-Einstein letters 1916-1955* (London: McMillan)