

# INTERVENTIONS AND CAUSALITY IN QUANTUM MECHANICS

Mauricio Suárez,  
Complutense University of Madrid,  
msuarez@filos.ucm.es

Forthcoming in *Erkenntnis*

Abstract: I argue that the Causal Markov Condition (CMC) is in principle applicable to the Einstein-Podolsky-Rosen (EPR) correlations. This is in line with my defence in the past of the applicability of the Principle of Common Cause to quantum mechanics. I first review a contrary claim by Dan Hausman and Jim Woodward, who endeavour to preserve the CMC against a possible counterexample by asserting that the conditions for the application of the CMC are not met in the EPR experiment. In their view the CMC is inapplicable to the EPR correlations – i.e. it neither obtains nor fails. The view is grounded upon the non-separability of the quantum state, and the consequent unavailability of interventions. I urge that whether interventions are available in EPR – and why – is a complex and contextual question that does not have a unique or uniform answer. Instead, I argue that different combinations of causal hypotheses under test, and different interpretations of quantum mechanics, will yield different answers to the question.

## 1. The Principle of Common Cause and the Causal Markov Condition

Hans Reichenbach first expressed the *Principle of the Common Cause* (PCC) as a statement regarding the existence of causes underlying what he called ‘improbable coincidences’. As part of his effort to define the direction of time in relation with that of open conjunctive forks, Reichenbach first of all asserted that: “if an improbable coincidence has occurred, there must be a common cause”. Without any further stipulation, and without any definition of a common cause, this statement remains essentially a postulate. It stipulates the existence of such a cause when a coincidence between two token events has occurred that is improbable in the sense that it is both lawlike and unexpected (i.e. it arises out of a correlation between the types, and it does not result from any causal relation between the tokens), then there must be some underlying event that causally explains away the coincidence. Let us refer to this general statement, which forms the first part of the PCC, as the *postulate of the common cause* (Suárez, 2007).

The PCC is not exhausted by the *postulate*, however, but it is typically combined with a specific *criterion of the common cause* providing a characterisation of common causes – and an essentially methodological as opposed to metaphysical guide to causal inference from statistics. Reichenbach’s own criterion was the very well known screening off condition, which took common causes to necessarily screen off their joint effects from each other. We say that  $c$  screens off  $a$  from  $b$  if and only if:  $P(a/b \& c) = P(a/c)$ .

The postulate and the criterion together make up the PCC. The order of introduction is as follows: the postulate is a requirement on the PCC, while the

criterion could in principle be replaced with an alternative. However, from a logical point of view, these two principles are independent. The PCC has been the subject of a great amount of debate and discussion in the literature, which I will not review here (but see San Pedro and Suárez, 2009, for a review). The only relevant issue for my purposes here surrounds determinism. The PCC is methodologically valid in the sense that it constitutes a guide for inference to those common causes that screen off. But as a guide it is always fallible because not all common causes screen off. In particular it has been argued to hold for *pseudo-deterministic* common causes, but not for all indeterministic common causes. Since it is impossible to tell in advance, by considering simply the statistical effects, which causes do as a matter of fact screen off, we are left with what is at best a heuristically valid method for inference.

In the last decade, however, the PCC has been superseded by another principle of inference, applicable to directed acyclic graphs, the so-called *Causal Markov Condition* (CMC). Yet, in the context of causal inference for EPR correlations, given some assumptions, the CMC does not differ substantially from the original PCC. It will not then come as a surprise that very similar issues over the validity of the principle of inference for pseudo-deterministic versus genuinely indeterministic causes arise in this new domain.

A directed acyclic graph is a representation of causal variables, where the different causal variables are represented as points and the causal relations as arrows, and where the arrows' starting and end points differ – thus containing and displaying no loops. A variable is then said to have descendents or parents if it is linked to other variables by means of such pointed arrows. Thus in the simplest graph between two variables,  $a \rightarrow b$ ,  $b$  is said to be a descendent of  $a$ , and  $a$  is said to be a parent of  $b$ . The CMC then stipulates the statistical relation that must obtain between parents and their joint descendents within a directed causal graph. In particular, it tells us that for any distinct pair  $(X, Y)$  in a set  $V$  of causal variables (i.e. those that are represented in a directed causal graph): if  $X$  does not cause  $Y$  then  $\text{Prob}(X / Y \ \& \ \text{Par}(X)) = \text{Prob}(X / \text{Par}(X))$ , where  $\text{Par}(X)$  is the set of all direct causes (parents) of  $X$  in  $V$ . Thus in a set with just three variables  $a$ ,  $b$  and  $c$ , where  $c$  is the only cause of  $a$ , and  $a$  does not cause  $b$ , CMC entails that:  $\text{Prob}(a / b \ \& \ c) = \text{Prob}(a / c)$ . This is true therefore for a common cause  $c$  of two correlated variables  $a$  and  $b$ , and it explains why the CMC entails the PCC (including the *criterion of the common cause*) in a three variable set with the causal structure typical of a common cause.

But why should any such principle be true? Why should common causes screen off (as the PCC states)? Why should the parents of a variable  $x$  screen off those variables that are not caused by  $x$  (as stated by the CMC)? Reichenbach seems to have taken the *criterion of the common cause* to express some a priori conceptual truth about causes, and conjoined this statistical characterisation with another conceptual truth about causes in terms of the mark-transmission criterion for those processes that genuine physical causes set in motion. Advocates of the CMC, by contrast, have defended it as an empirical and contingent truth at best – one which needs some backing, or grounding upon, further principles. They have in particular appealed to the *modularity* assumption as a way to ground the CMC

(Hausman and Woodward, 1999, p. 545). Roughly stated, modularity claims that the equations that relate the values of the different variables in a causal set  $V$  are invariant under changes of values of some of its members. In other words modularity expresses a form of robustness of the equations that describe the functional relations between the variables in a causal set. More precisely, a set is only modular if there exists a range of values of some of its members under which the equations relating all other variables in  $V$  remain invariant. Hausman and Woodward explicitly link this idea to the notion of *intervention*, as follows:

**Modularity (MOD):** For all subsets  $Z$  of the variable set  $V$ , there is some non-empty range  $R$  of values of members of  $Z$  such that if one intervenes and sets the value of the members of  $Z$  within  $R$ , then all equations except those with a member of  $Z$  as a dependent variable (if there is one) remain invariant.

There are a number of problematic assumptions and issues in relation with this condition. Some are well known and refer to the relation between MOD and the CMC. Although MOD is supposed to ground CMC, the derivation of the CMC from MOD is by no means trivial. There are three worries, which we will essentially bracket out in this paper, but deserve mention.

First, the proof of CMC out of MOD requires further substantial assumptions, such as the need for correlations to be explained, causal sufficiency, and the presence of unrepresented causes. These assumptions are here described for completeness, even though I have nothing substantial to say about them – and will simply assume them for the sake of our argument. The first assumption is the equivalent of what was referred to above as the *postulate of the common cause*, namely the claim that every law-like correlation has some causal explanation, and which Hausman and Woodward (1999, p. 524) refer to as **CM1**: “If  $X$  and  $Y$  are probabilistically dependent, then either  $X$  causes  $Y$ , or  $Y$  causes  $X$ , or  $X$  and  $Y$  are effects of some common cause  $Z$  in the set of variables  $V$ ”. As in the case of the postulate of the common cause, CM1 is a requirement on the Causal Markov Condition, yet it is logically independent from any characterisation of common causes. The second assumption (‘causal sufficiency’) requires that all common causes of variables in the causal set  $V$  must be included in that set. Hence if  $X$  and  $Y$  are included in set  $V$  then any common cause  $C$  of both must also be included. The third assumption, however, states that there are unrepresented causes, i.e. that there are causes of variables in the set  $V$  that are not included in that set. Given causal sufficiency, such causes can only be partial single causes of particular variables in  $V$ , or their values, often referred to as “exogenous” or “independent” variables. In the present context, as we shall soon see, ‘interventions’ are amongst such unrepresented causes. The first worry regarding CMC is of course that these three assumptions are all rather substantial, and may be coherently denied. However, I shall not enter the discussion regarding their status, since it is besides my purposes here.

The second worry regarding CMC is that, even granted the above three assumptions, the proof of CMC out of MOD has been contested (Cartwright, 2002). Of course, if CMC does not follow from MOD, regardless of additional assumptions, it cannot be grounded upon the intuitions underlying MOD, in any case. The

difficulties have been debated between Hausman and Woodward on the one hand, and Nancy Cartwright on the other, and I will not rehearse them here. Instead I shall simply assume, contra Cartwright, that Hausman and Woodward's proof holds. From the point of view of the argument this concedes no significant ground, since if the proof did not go through, this would not mean that CMC does not apply to EPR, as Hausman and Woodward claim. It would rather mean that in order to apply it there is not even a need to consider the nature of interventions in EPR. There is thus a sense in which the route taken here to prove the applicability of CMC is the hardest possible route – consisting in assuming the whole of Hausman and Woodward's theory of causal inference and then showing that, contrary to their claim, it may well apply to quantum mechanics.

The third and final worry is that the link between MOD and interventions is subtle, at best, since there seems to be *prima facie* no reason why the range of values of the members of Z referred to in the definition above should in any way be achievable by means of an "intervention" – however we characterise the term. What seems to rather be essential to the statement of modularity is that such a range of values exists regardless of how they come about, or indeed whether they come about at all. Again I will not enter this issue in detail here, but will just hope that, as the logic of our argument develops, the reader will realise that nothing substantial hinges upon any of these assumptions. My objections to Hausman and Woodward's claims regarding the applicability of CMC to quantum mechanics do not rely on any failure of the presumed general link between interventions and the CMC. On the contrary, I want to resist Hausman and Woodward's conclusions regarding quantum mechanics anyway, for independent reasons, which are unrelated to their proof of CMC out of MOD. Hence it will be assumed in this paper that the proof goes through and that interventions are required for modularity precisely as described by Hausman and Woodward.

## 2. Is the Causal Markov Condition inapplicable to Quantum Phenomena?

Note that the question above is not whether the CMC is violated in quantum mechanics. If it were, then quantum mechanics would show the CMC to be false. And indeed a respectable tradition takes it that the EPR correlations constitute a counterexample to the PCC (and therefore to CMC), which is thus falsified by quantum phenomena (for an influential argument to this effect, see Van Fraassen, 1982). The advocates of the CMC argue by contrast that the CMC is not falsified; rather it is inapplicable in this domain: the conditions for its application (positive or negative) just do not obtain. In particular, Hausman (1999) and Hausman and Woodward (1999, pp. 564-567) both argue that modularity fails in the EPR experiment because there are no interventions available in this context. They provide two different reasons for this claim, but both turn out to be related to the non-separability of the quantum state. I will review these reasons briefly. But first note that the overall structure of Hausman and Woodward's argument is rather peculiar and seems to boil down to the following sequence of premises:

i) There are no possible interventions in an EPR experiment. Therefore:

ii) MOD does not apply to the EPR correlations. Thus:

iii) CMC does not apply to the EPR correlations. Therefore:

iv) The EPR correlations are no counterexample to the interventionist (or manipulability) account of causality.

This paper is mainly concerned with questioning premise i). It is obvious that if i) is not generally true, then there is no reason for ii) to be true. Moreover, even if i) were true, ii) does not necessarily follow – for, as was mentioned in the previous section, there is no clear reason why interventions need to be appealed to for MOD in general to hold with respect to different values of the variables involved. Finally, let it be mentioned that even if premises i) and ii) were provably true, iii) would not necessarily follow either. CMC follows from MOD only provided the additional assumptions of explaining correlations, causal sufficiency, and unrepresented causes. So the failure of MOD does not impugn CMC, even granting those assumptions.<sup>1</sup> Finally, does the failure of CMC to apply in this case save the manipulability theory of causation (i.e. the view defended in Woodward, 2003) from refutation? This depends very much on what the manipulability theory precisely asserts about the availability of interventions, in practice and in theory. Let us then turn to the concept of intervention employed in the manipulability theory of causation.

According to Woodward (2003, p. 98), I is an *intervention variable* for X with respect to Y if and only if it meets the following four conditions, which I describe as sub-indexes to I for “Intervention”:

I<sub>1</sub>: I causes X

I<sub>2</sub>: I acts as a switch for all other variables that cause X (within a certain range of values of X).

I<sub>3</sub>: Any directed path from I to Y goes through X.

I<sub>4</sub>: I is statistically independent of any variable Z that causes Y and that is on a directed path that does not go through X.

An *intervention* (on X with respect to its putative effect Y, where X and Y are both included in the corresponding variable set V) is then just the setting of a value of an intervention variable for X with respect to Y. The central idea is that we intervene to set the value of X in order to figure out how robust the link between X and Y is with respect to changes in values of X. We may model an intervention by means of a graph where the pointed arrows stand for directed causes, while the

---

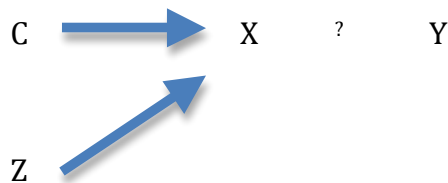
<sup>1</sup> Hausman and Woodward (1999, p. 553) argue further that, given the three background assumptions mentioned, MOD holds if and only if CMC holds, i.e. they are logically equivalent conditional on the background assumptions. If so the step from ii) to iii) in the above argument does hold, but only conditional on the assumptions.

question mark represents the fact that we are unsure as to the nature of the relation between X and Y:



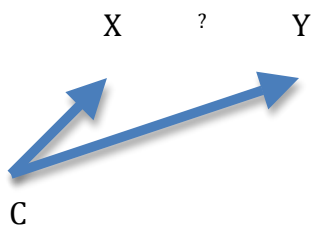
Let us get clear on what each of these clauses rules out. The first clause  $I_1$  rules out any intervention that is not a cause of X. It may seem circular to include the requirement that interventions be causes, since interventions are introduced precisely with the aim of testing causal relations (we intervene on X by means of I in order to find out whether X is indeed a cause of Y). However, the manipulability theory is not meant to analyse causation, so the circularity is unthreatening. The language of interventions is causal, and it seems reasonable to suppose that an intervention on a variable is itself a cause of the value of the variable.

The second clause  $I_2$  rules out any cause of X, whether they are or not represented in the set  $V$ , that fails to switch off all other causes of X. In other words, I is not an intervention variable unless it is the only cause acting upon X and suffices to determine the value of X. In particular, consider the following graph where the arrows represent directed causal relations, while the question mark expresses the fact that we are unsure as to the nature of the relation between the X and Y variables. In this causal structure the variable C is not an intervention upon X with respect to Y because there is another cause of X, namely Z, which contributes to fixing its value:



It then follows that acting on X by means of C cannot fix the value of X and is therefore unhelpful as a test of the causal relation between X and Y. For C to be a proper intervention I, it would have to switch off Z as well as any other variable that can causally contribute to the value of X.

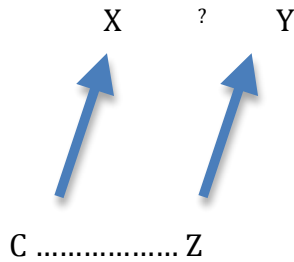
Let us now consider the third clause  $I_3$ . This clause rules out any cause C of X that may causally affect Y via a path that does not include X. In particular it rules out a cause C of X that is also a direct cause of Y, as in the graph below:



In the above structure C is not an intervention on X with respect to Y even though it fixes the value of X because it also fixes the value of Y, independently of the value of X. Hence it is unhelpful in order to assess the dependency of Y on X and

whether or not the functional relation between X and Y is invariant with respect to the values of the cause (within a certain range). For C to be an intervention it would have to act on Y exclusively via its effect upon X.

Finally the fourth clause  $I_4$  rules out causes of X that are statistically correlated with other causes of Y, regardless of whether or not they are causally related. Hence in the graph above even if there were no arrow pointing from C to Y, C would not be an intervention on X with respect to Y if it turns out to be statistically correlated with Z, which is a distinct cause of Y:



In the graph above the statistical correlation between C and Z is represented by a discontinuous line. The correlation may or may not have a causal source (i.e. there may or not be a common cause of C and Z); the existence of the correlation is enough on its own to invalidate the claim that C is an intervention on X (with respect to Y). This excludes any correlations with any causal variables that lie in the causal path from C and Y via X (in particular it excludes any intermediate causes lying between X and Y).

Hausman and Woodward’s claim that there are no possible interventions in EPR therefore boils down to the claim that there is no variable that satisfies all the defining clauses of an intervention. There are instead at best some C variables that fail to fulfil at least some among the four conditions, as in the graphs displayed above. We may now proceed to analyse this claim in detail. First, in section 3 I briefly recapitulate the debate over causality in quantum mechanics, and describe the nature of the EPR correlations. Then, in section 4, I describe and criticise Hausman and Woodward’s analysis.

### 3. Quantum Correlations

At its inception quantum mechanics was often presented as a non-causal theory (Von Neumann, 1955). The founding parents often identified causality and determinism (in other words they held what we may call, following Suárez and San Pedro (2011) the *c=d identity*), and therefore understood the emergence of the new irreducibly indeterministic quantum mechanics as an indictment of causality. Even the opponents of the new quantum mechanics shared the *c=d identity*. Thus, for instance, David Bohm showed that von Neumann’s theorem against hidden variables involves essentially questionable premises, thus paving the way for hidden variables. But while Bohm and Von Neumann disagreed regarding the status of causation in quantum mechanics, they agreed that the fortunes of causation and determinism were essentially linked. Bohmian mechanics restores

determinism to the description of the most fundamental dynamical properties of quantum objects, and it allows for the ascription of continuous trajectories to particles. It is sometimes known as the causal theory, but it is strictly speaking distinct from orthodox quantum mechanics in the deterministic character of its equations of motion. I shall return to Bohmian mechanics in due course in the final section. Let me just note now that orthodox quantum mechanics and Bohmian mechanics are just two amongst many other interpretations or theoretical frameworks that may be applied to the EPR experiment. Each provides a particular physical model of the experiment and the ensuing correlations. And with the advent of probabilistic causation, the fates of determinism and causality are of course no longer tied. Thus some interpretations provide straightforward causal models, others not so obviously so; but, at any rate, the indeterminism of the orthodox interpretation is no longer a reason to rule out causal models for the EPR correlations.

The bare, uncontested, and interpretation-free facts regarding the EPR experiment may be summarised as follows. The thought-experiment was first introduced in a famous article published in 1935 (Einstein, Podolsky and Rosen, 1935). Later on, David Bohm provided the version that is almost universally adopted these days, in which two entangled electrons A and B are emitted from a source in opposite directions. The spin component of each electron can be measured as each electron travels towards a measurement device in its respective wing – the details of the measurement procedure need not detain us here. The values of the spin variable of each electron in the singlet state may be either ‘spin-up’ ( $\uparrow$ ) or ‘spin-down’ ( $\downarrow$ ) with probability  $\frac{1}{2}$ . Let us then denote the measurements as  $a$  on particle A, and  $b$  on particle B. And let us denote the corresponding measurement outcome events on each particle as  $\uparrow_a$ ,  $\downarrow_a$ ,  $\uparrow_b$ , and  $\downarrow_b$ . In other words, each measurement may be represented without ambiguity as  $a$ ,  $b$ ; while each measurement outcome event will be  $\uparrow_a$ ,  $\downarrow_a$ ,  $\uparrow_b$ , or  $\downarrow_b$ . The state of the entangled pair, the singlet state, may then be expressed as follows:

$$\psi = \frac{1}{\sqrt{2}}(|\uparrow_a\rangle|\downarrow_b\rangle - |\downarrow_a\rangle|\uparrow_b\rangle).$$

The measurement events at each wing of the experiment, such as  $\uparrow_a$ , and  $\downarrow_b$ , are space-like separated events, i.e. they lie outside each other’s light cone. (Hence in the diagram of the experiment in Figure 1, there exists no time-like worldline from  $b$  to  $a$ , or vice versa).



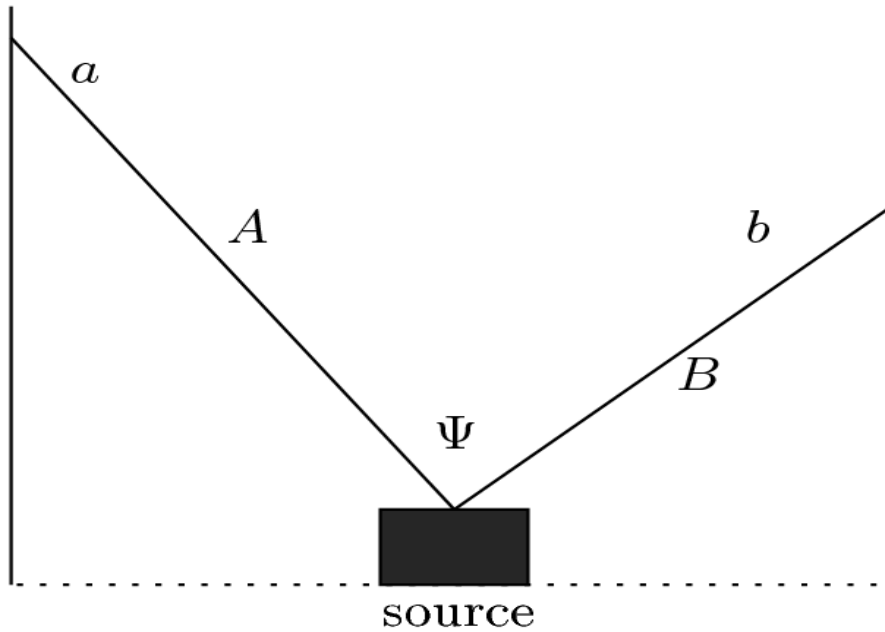


Figure 1: Spacetime representation of a typical EPR experiment.

The ‘EPR correlations’ are the conditional statistical dependencies exhibited by the outcomes on each wing in this experiment, as predicted by quantum mechanics for the singlet state. More specifically for measurements along the same direction of spin the outcomes are anti-correlated, as follows:

$$P(\uparrow_a / \downarrow_b) = P(\downarrow_b / \uparrow_a) = 1; \text{ while } P(\uparrow_a / \uparrow_b) = P(\downarrow_a / \downarrow_b) = 0.$$

The question for the PCC and the CMC is how to explain these correlations causally. Hausman and Woodward claim that these are brute nomological correlations that cannot be explained causally – and in particular they claim the methodology of causal inference they defend is inapplicable in this domain. In the next section I describe these claims, and assess them critically.

#### 4. Hausman and Woodward on CMC and the EPR Correlations.

Modularity (MOD) requires that there exist separate mechanisms for the causal relation between X and Y, and for any other mechanisms underlying Y. Thus an intervention on X does not disrupt these alternative mechanisms, and brings about a change in Y only as a function of its change in the value of X if and only if X and Y are causally related in the appropriate way. In other words in order MOD (and CMC) can only be employed to test causal relations in a domain that satisfies this minimal requirement (that the causal mechanisms be separable). Otherwise neither MOD nor CMC have genuine application, and the issue as to the causal relation between X and Y is left unanswered. Hausman and Woodward’s argument against the applicability of CMC to the EPR correlations is based upon the premise that the entangled particles in the EPR experiment precisely lack such separable mechanisms. As they write (Hausman and Woodward, 1999, p. 566): “The notion of an intervention with respect to one of the measurement events is not well-

defined in the EPR phenomena, because the distinction between intervening with respect to X and acting directly on both X and Y cannot be drawn.”

This argument, concluding for the unavailability of interventions in EPR, depends upon a particular interpretation of the quantum state as non-separable in a strong but rather mysterious ontological sense (Hausman and Woodward, 1999, p. 566): “In some way that is difficult to understand, the two particles constitute a single composite object, even though they may be at spacelike separation from each other”. It is obvious that if particles A and B are actually the ‘same’ particle, then there can be no intervention on the spin value  $b$  of particle B which is not also an intervention upon the spin value  $a$  of particle A, since both variables  $b$  and  $a$  are as matter of fact properties of the same system – they are therefore just a part of an overarching and complex causal mechanism.

However, the most precise approach to holism does not refer to non-separable objects, but rather to non-separable events – i.e. the particles are individuated by their distinct spatio-temporal location, but the measurement outcome events are non-separable in the sense that they are non-locally dependent upon each other. Indeed Hausman and Woodward main reference is an old paper by Brian Skyrms (1984) defending just this sort of non-locality for measurement outcome events.<sup>2</sup> Their reasoning is as follows: “The measurement result on one wing is not really a distinct event from the result on the other wing but rather both comprise a ‘single, indivisible non-local event’”. The term “single, indivisible non-local event’ can indeed be found in Skyrms, but not the inference to the non-separability of states that Hausman and Woodward draw from it: “For this reason, it is wrong to think of the measurement process performed on one particle as directly affecting only the state of that particle and affecting the other particle if at all only through the change it produces in the first particle” (ibid, p. 566).

There is indeed a significant difference between ascribing non-separable states to the particles, and – in strict accordance with Skyrms’ phrase above – describing the outcome events as non-separable. The former interpretation seems contradicted by the axiom of reduction, which makes it always possible ascribe separate states to each of the particles at any given time, by tracing out the degrees of freedom from the entangled state of the pair. Thus in the singlet state, both particles at any given time are assigned mixed states represented by the statistical operators  $W_A$  and  $W_B$  that result out of tracing out as follows:

$W_A = \frac{1}{2}(|\uparrow_a\rangle\langle\uparrow_a| \otimes |\downarrow_b\rangle\langle\downarrow_b|)$ , and  $W_B = \frac{1}{2}(|\uparrow_b\rangle\langle\uparrow_b| \otimes |\downarrow_a\rangle\langle\downarrow_a|)$ . There is nothing particularly obscure or mysterious about these states, other than the fact that they together do not describe all the information that is relevant to the relations between the particles’ properties. As Schrödinger once famously put it, some of the relevant information (precisely that which regards the conditional statistical dependencies between the measurement outcomes) is ‘squandered into

---

<sup>2</sup> Skyrms (1984, p. 255) attributes the view to David Lewis, but it was a rather entrenched view at the time – cf. Michael Redhead’s very similar ‘ontological locality’ or OLOC (Redhead, 1987).

conditional statements that operate between the subsystems' (Schrödinger, 1935, p 161).

How about the latter reading, which is more strictly in accordance with Skyrms' original phrase? What does it mean to say that the measurement outcome events are in fact just the same 'single, indivisible, non-local event'? For our purposes there is no better way to paraphrase this statement than precisely in terms of the manipulability account of causation. We merely need a translation from the language of events as causal relata, to that of variables, which can be achieved by means of bi-valued 'yes-no' variables. Then Skyrms' phrase may be interpreted to simply mean that there is no way to intervene upon the first event-variable independently of the second. Any putative intervention upon  $b$  must ipso facto also be understood to be an intervention upon  $a$ , and questions relative to the strict dependence of the outcome  $a$  upon the outcome  $b$  cannot arise (in the terms in which the manipulability theory would formulate such questions).

The view is coherent, but is it true? Can there be no proper interventions upon  $a$  – and why not? Hausman and Woodward presuppose that there can only be one answer to this question, and they argue that it is negative. As they put it in their more recent paper (Hausman and Woodward, 2004, page 157, fn. 5): "Recall [...] that examples from microphysics [...] are cases in which no arrow-breaking interventions on the allegedly distinct effects are possible, which is to say that only a single mechanism is involved." To be sure, measurements cannot be interventions that determine specific values for observables. It is well known that the spherical symmetry of the singlet state entails that at most we can predict the outcome of  $b$  (i.e. whether the outcome is spin up or spin down) with probability  $\frac{1}{2}$ , regardless of what direction of spin we measure. However, Hausman and Woodward accept indeterministic causes that may only fix the probability of their effect (Hausman and Woodward, 1999, p. 570; see also Woodward, 2003). Now, certainly, the measurement itself is not something that an agent can bring about – it is rather simply the interaction of the particle with the measurement device at rest in the laboratory frame. The setting of the measurement device, however, is something an agent definitely can bring about; and in fact an agent is required to set the device to measure spin along a particular direction of spin. The question for us then is whether such setting events (let us refer to them as  $S_a$  and  $S_b$ ) satisfy the four clauses on interventions, and can consequently be said to be proper interventions. I believe that the answer to this question is not in any way as obvious as Hausman and Woodward presuppose.

For instance, can  $S_b$  be said to be a cause of  $b$  (clause  $I_1$ )? The answer to the question is surprisingly elusive. If by 'being a cause' we mean 'actively changing the value' of  $b$ , or its probability, then the answer must be negative, for that probability is never anything other than  $\frac{1}{2}$  (if it is defined at all). But if we just mean 'determining the value' of  $b$ , or its probability, then certainly the setting event determines that the probability of a particular outcome of the spin measurement is  $\frac{1}{2}$ . True, it does not determine it uniquely, since one can choose to measure along any other axis with the same outcome probability. But nowhere in the manipulability theory is there a requirement for causes to determine the

values of their effects uniquely – and it would be rather counterintuitive to insist on such a thing anyway, independently of the theory of causation espoused.

Now, let us continue, can  $S_b$  said to act as a switch of all other variables that cause  $b$  (clause  $I_2$ )? There is certainly an issue regarding whether it makes sense to suppose that there are any other variables in the hypothetical causal structure that causes  $b$ . Different interpretations of quantum mechanics may differ on this point. For instance, in the orthodox interpretation of quantum mechanics, there are no hidden variables, and the causal structure is complete. So there is no question of switching off any further variables. Other theories, such as Bohmian mechanics, postulate further structure, so the question does meaningfully arise. Yet, those further variables seem to have no effect whatsoever upon the outcome of  $b$ , or its probability, which given the setting is  $\frac{1}{2}$  on any model. So, I don't see any reason to suppose that  $S_b$  would not generally comply with clause  $I_2$ , regardless of whether there are further causes of  $b$ . And in particular if  $b$  has no other causes, as the hypothesis of a direct link between  $b$  and  $a$  implicitly suggests under the orthodox interpretation, then clause  $I_2$  is automatically satisfied without any further ado.

Is there any direct path from  $S_b$  to  $a$  not going through  $b$  (clause  $I_3$ )? I see no reason in principle to suppose so, on the bare description of the causal hypothesis that we are entertaining (i.e. just that  $b$  is a cause of  $a$ ). This is particularly evident since the existence of such a direct path would seem to violate special relativity. Or at least, this view is in the background of the common argument for why violations of the so-called parameter independence condition contradict the special theory of relativity.<sup>3</sup> The argument presupposes that the violation of parameter independence (PI) is due to a causal connection or influence that can transmit a signal from the setting event on the one wing to the outcome event on the distant wing. By contrast, it is assumed that the violation of outcome independence (OI) is not due to such a causal connection, and cannot be employed to transmit information or signals, because the outcome events are not controllable. It follows from the assumption that  $b$  may cause  $a$  in the sense prescribed without any violation of special relativity, but  $S_b$  may not cause  $a$  in the same sense, on pain of a violation of special relativity, since the setting event  $S_b$  is controllable.

Now, note that I am not here endorsing the argument above against violations of parameter independence, but only the causal presupposition that underlies it. This is important since the analysis of the failure of these two conditions that I have just given above is certainly not conclusive. Although it is

---

<sup>3</sup> See Jarrett (1984) for the original argument. Parameter independence, stated in a simplified form, asserts the statistical independence of the far away outcome on the nearby setting, given the state and far away setting:

$\Pr(a/S_b \& S_a \& \psi) = \Pr(a/S_a \& \psi)$ . Outcome independence asserts the statistical independence of the far away outcome on the nearby outcome, given the state and both settings:  $\Pr(a/b \& S_b \& S_a \& \psi) = \Pr(a/S_b \& S_a \& \psi)$ . The interpretation of their failure in causal terms that I provide in the text is standard amongst physicists and many philosophers of physics. Yet other philosophers, and experts on causal inference, do not accept that a causal or physical dependence typically has such a straightforward representation in the statistics.

common amongst physicists and philosophers of physics, the analysis has been contested – and rightly so in my view. In particular, many philosophers and experts on causal inference would not accept that a causal or physical dependence typically has a straightforward representation in the statistics. The statistical dependencies that violations of (PI) and (OI) express may mask complex causal relations, and need not indicate any direct causal connections between the conditioned upon and the conditional events. Thus, Maudlin (1994) argues that violations of parameter and outcome independence stand very much on identical grounds vis a vis special relativity: Whether a violation of a statistical condition, such as either (PI) or (OI), contradicts relativity has to do with the underlying causal structure and not with the violation of the statistical condition itself. It is important to emphasise that one can agree with this criticism of the standard argument against violations of (PI) without disagreeing with the causal presupposition that underlies it – namely that  $b$  and not  $S_b$  can be a cause of  $a$  without violation of special relativity because it is not controllable in the way that  $S_b$  is. The causal presupposition may be correct even if its representation as a statistical condition is not. And only the underlying presupposition is required to show that clause  $I_3$  on interventions indeed applies to  $S_b$  as an intervention on  $b$  with respect to  $a$ .

Yet, it has to be emphasised too that the above analysis in favour of  $I_3$  presupposes the standard or orthodox interpretation of quantum mechanics. For instance, on a causal interpretation of Bohmian mechanics, the quantum potential may act as an intermediate cause between  $S_b$  and  $a$  – one that moreover does not in any way go through  $b$  (see Holland, 1993, Ch. 11). On any other interpretation of quantum mechanics (including, arguably, the ‘minimal’ interpretation of Bohmian mechanics) such an assumption seems far-fetched if not outright impossible. So, on one interpretation  $S_b$  may fail to satisfy clause  $I_3$  on interventions (because there is a different causal structure underlying the correlation), while on a different interpretation it may fully satisfy it. This suggests the more general thought that whether a particular variable is or is not an intervention in the EPR experiment depends not just on the particular causal hypothesis under test, but also very sensitively on the overall interpretation or version of quantum mechanics that is adopted. It also makes it clear that, contrary to what Hausman and Woodward presuppose, no precise assessment of the applicability of interventions in EPR can be drawn in general – independently of both the causal hypothesis and the interpretation adopted.

Let us finally consider clause  $I_4$ . The question is whether  $S_b$  may be statistically correlated with any of the causes of  $a$ . In the causal hypothesis that we are exploring, the only cause of  $a$  other than  $b$  is  $S_a$ . Are  $S_a$  and  $S_b$  statistically correlated? It may appear so from the spherical symmetry of the singlet state since it presupposes the outcome events to be perfectly anti-correlated for measurements of spin along the same direction. But as a matter of fact this is just a peculiarity of our representation. There is no reason in principle why the experimenters on both wings must set their measurement devices to measure along the same direction of spin. Once  $b$  is fixed,  $a$  may be fixed in any chosen direction, and there is no reason in principle to expect any correlation between these settings. Clearly, if  $a$  and  $b$  do not measure along the same directions, the

singlet state will not issue anti-correlated outcomes. The predictions however are well defined in quantum mechanics, and were fully described and taken into account in the derivation of the complete Bell inequalities (see e.g. Redhead, 1987, Ch. 3). The values of  $a$  and  $b$  are freely set by the experimenters so there is no reason to expect them to be correlated.<sup>4</sup>

Thus the only reason to suspect, on the orthodox interpretation of quantum mechanics, that there may be no interventions available in the EPR case is a failure of clause  $I_1$ . I have argued above that on a natural reading of what it is for a setting event on a measurement device to be a cause of a measurement outcome event in that very device, there appears to be no issue in reconciling the claim that, for example,  $S_b$  is a cause of  $b$ . If so, there appears to be no real issue in the claim that the setting events are interventions in the full sense of the term that allows us to test the claim that the measurement outcome events  $a$  and  $b$  are causally related.

This analysis, however, just applies to one particular and very simple causal hypothesis, namely that  $b$  is a cause of  $a$ . We could also invoke more complex causal hypotheses where the notion of intervention may or may not turn out to be applicable. There are also sensitive issues regarding the interpretation of quantum mechanics that is adopted. Thus I argued above that if the EPR experiment is given a Bohmian model with a ‘causal’ interpretation, clause  $I_3$  turns out false – although a different causal hypothesis is true.<sup>5</sup> So for the very same causal hypothesis, the claim that  $S_b$  is an intervention depends upon the physical model and interpretation of EPR that are adopted.

## 5. Conclusions

The claim that EPR is no counterexample to the CMC, because there are no interventions available in the quantum domain, has been contested. It has been argued on the contrary that the question: “is EPR indeed a counterexample to the CMC?” is still very much a valid and pertinent question. The question has often been asked in the past, and received a negative answer (Van Fraassen, 1982; although see also Rédei et al., forthcoming, for a different formulation that rejects the received answer). Yet, any attempt to settle it in the abstract is bound to disappoint. The question requires a framework, in the form of both a precise causal hypothesis to test, and a specific interpretation of quantum mechanics. There is no way to settle such matters in the absence of detailed information regarding both critical aspects. And there is no reason at all to expect a uniform answer across all such frameworks, and quite a lot of evidence to suggest to the contrary that the answers to the question will differ greatly depending on frameworks.

---

<sup>4</sup> And indeed, as a referee pointed out, in a typical EPR experiment the settings are, as a matter of fact, chosen randomly, so as to expressly be uncorrelated.

<sup>5</sup> This further causal hypothesis, involving the quantum potential is moreover perfectly in line with the CMC (see Suárez and San Pedro, 2011, pp 187-190).

## ACKNOWLEDGEMENTS

I would like to thank: the editors of the volume for their careful editorial work; the organisers of the CaEitS conference at Ghent University (Erik Weber and Bert Leuridan) for their encouragement; the audience at Ghent for their reactions; two anonymous referees of *Erkenntnis* for their acute and constructive suggestions; my graduate students and the members of the LSE causality reading group for comments; and Iñaki San Pedro, in particular, for many conversations on related topics over the years – as well as the figure of the EPR experiment (reprinted from Suárez and San Pedro, 2011, with Springer’s permission). Financial support is acknowledged from the Spanish Government (grants FFI2008-06418-C03-01 and FFI2011-29834-C03-01).

## REFERENCES

- Cartwright, N. (2002), “Against Modularity, the Causal Markov Condition, and any Link between the two: Comments on Hausman and Woodward”, *British Journal for the Philosophy of Science*, 53, pp. 411-453.
- Einstein, A., B. Podolsky, and N. Rosen (1935), “Can Quantum Mechanical Description of the World Be Considered Complete?”, *Physical Review*, 47, pp. 777-780.
- Hausman, D. (1999), “Lessons from Quantum Mechanics”, *Synthese*, 121, pp. 79-92.
- Hausman, D. and J. Woodward (1999), “Independence, Invariance and the Causal Markov Condition”, *British Journal for the Philosophy of Science*, 50, pp. 521-583.
- Hausman, D. and J. Woodward (2004), “Modularity and the Causal Markov Condition: A Restatement”, *British Journal for the Philosophy of Science*, 55, pp. 147-161.
- Holland, P. (1993), *The Quantum Theory of Motion: An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics*, New York: Cambridge University Press.
- Jarrett, J. (1984), “On the Physical Significance of the Locality Conditions in the Bell Arguments”, *Nous*, 18, pp. 569-589.
- Maudlin, T. (1994), *Quantum Non-locality and Relativity*, Oxford: Blackwells.
- Rédei, M. et al. (Forthcoming), *The Principle of the Common Cause*, Cambridge: Cambridge University Press.

- Redhead, M. (1987), *Incompleteness, Non-locality and Realism*, Cambridge: Cambridge University Press.
- Skyrms, B. (1984), "EPR: Lessons for Metaphysics", *Midwest Studies in Philosophy*, 9, pp. 245-255.
- San Pedro, I. and Suárez, M. (2009), "The Principle of Common Cause and Indeterminism: A Review" in González Recio, J. L. (ed.), *Philosophical Essays on Physics and Biology*, Georg Olms, pp. 223-250.
- Schrödinger, E. (1935), "The Present Situation in Quantum Mechanics", a translation of Schrödinger's cat paradox paper reprinted in Wheeler and Zurek (eds.), *Quantum Theory and Measurement*, Princeton: Princeton University Press.
- Suárez, M. (2007), "Causal inference in quantum mechanics: A reassessment", in F. Russo and J. Williamson (eds.), *Causality and Probability in the Sciences*, London: College Texts, pp. 65-106.
- Suárez, M. and I. San Pedro (2011), "Causal Markov, Robustness and the Quantum Correlations", in M. Suárez (ed.), *Probabilities, Causes and Propensities in Physics*, Synthese Library, 347, Dordrecht: Springer, pp. 173-193.
- Van Fraassen, B. (1982), "The Charybdis of Realism: Epistemological Implications of Bell's Inequality", *Synthese*, 52, pp. 25-38. Reprinted with corrections in J. Cushing and E. McMullin (eds.), *Philosophical Consequences of Quantum Theory*, Notre Dame: Notre Dame University Press, 1989.
- Von Neumann, J. (1955), *Mathematical Foundations of Quantum Mechanics*. Princeton: Princeton University Press. First published in German in 1932 as *Mathematische Grundlagen der Quantenmechanik*, Berlin: Springer.
- Woodward, J. (2003), *Making Things Happen: A Theory of Causal Explanation*, Oxford: Oxford University Press.