Appeared in M. Suárez and A. Bird (eds.), Dispositions, Causes and Propensities in Science, *Theoria, vol. 19, 3, September 2004, pp. 271-300.*

CAUSAL PROCESSES AND PROPENSITIES IN QUANTUM MECHANICS

Mauricio Suárez (Universidad Complutense de Madrid)

Abstract:

In an influential article published in 1982, Bas Van Fraassen developed an argument against causal realism on the basis of an analysis of the Einstein-Podolsky-Rosen correlations of quantum mechanics. Several philosophers of science and experts in causal inference – including some causal realists like Wesley Salmon – have accepted Van Fraassen's argument, interpreting it as a proof that the quantum correlations cannot be given any causal model. In this paper I argue that Van Fraassen's article can also be interpreted as a good guide to the different causal models available for the EPR correlations, and their relative virtues. These models in turn give us insight into some of the unusual features that quantum propensities might have.

Keywords: Philosophy of Science, Quantum Mechanics, Causality, Propensities, Realism.

1. The Einstein-Podolsky-Rosen (EPR) Correlations

In 1935 Einstein, Podolsky and Rosen published their notorious paper [1] describing a thought experiment with correlated entangled pairs of particles. The version of this experiment that has been most discussed, and gives rise to the real experiments that have actually been carried out, is due to David Bohm.[2]

In this Einstein-Podolsky-Rosen-Bohm experiment two particles ("1" and "2") with entangled states (because they have interacted in their past, or because they have been created simultaneously in the same decay event), move in opposite directions. In a Minkowski space-time diagram, both particles describe symmetric paths along the time axis (see figure 1). The initial angular momentum is zero, so their values of spin must be correlated throughout. Any particle's spin can be measured by means of a Stern-Gerlach apparatus. (Essentially a magnetometer that impresses a force upon the particle proportional to its spin value, thereby correlating perfectly the particle's position with its spin value at the time the particle interacts with the magnetometer. A Stern-Gerlach apparatus can be rotated along 360 degrees, in order to measure the particle's spin value along any direction).

[1] Einstein, Podolsky and Rosen, (1935).
[2] Bohm (1951, cap. 22).

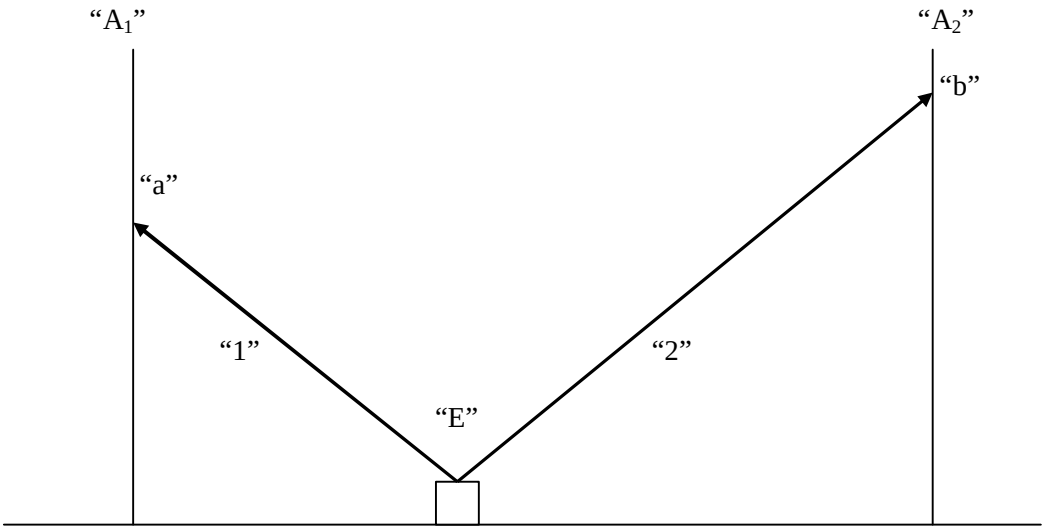


Figure 1.

Each particle's spin is measured in its corresponding wing of the experiment; the outcomes of each experiment are denoted by "s_1" and "s_2" – these are known as the "outcome-events". (Lines "A_1" y "A_2" represent the world lines of both measurement apparatuses, which are at rest with respect to the laboratory frame). An important feature of the EPR-Bohm experiments is that these outcome-events are spacelike connected, i.e. they lie outside each other's light-cone. Thus a signal from one event to the other must travel at speed greater than the speed of light, during a finite part of its trajectory at least. As will be pointed out later, the implications of this fact with regards the special theory of relativity are both deep and complex, and have been the object of an intense debate. [3]

According to quantum mechanics, there are only two possible values of spin in any direction of measurement (θ): positive spin ($\uparrow_\theta$) and negative spin ($\downarrow_\theta$). Quantum mechanics describes the spin state of the composite system of both particles at the time of emission and measurement by what is known as the singlet state:

$$\Psi = {}^{1}\!/_{\sqrt{2}}\left(\left|\uparrow_{1q}\right\rangle\left|\downarrow_{2q}\right\rangle - \left|\downarrow_{1q}\right\rangle\left|\uparrow_{2q}\right\rangle\right).$$

(Here "1" and "2" refer to each particle.) The theory offers two kinds of probabilistic predictions. First, it offers predictions about the outcomes of measurements performed on each particle. To calculate these, we must first apply what is known as the axiom

[3] See e.g. Maudlin (1995).

of reduction, which allows us to derive the state of each particle, individually taken, consistent with their being in the composite singlet state: [4]

$$W_1 = \frac{1}{2}\left|\uparrow_{1q}\right\rangle\left\langle\uparrow_{1q}\right| + \frac{1}{2}\left|\downarrow_{1q}\right\rangle\left\langle\downarrow_{1q}\right|,$$
$$W_2 = \frac{1}{2}\left|\uparrow_{2q}\right\rangle\left\langle\uparrow_{2q}\right| + \frac{1}{2}\left|\downarrow_{2q}\right\rangle\left\langle\downarrow_{2q}\right|.$$

Using these two states quantum mechanics allows us to make two types of predictions:

1. Probabilities of outcomes of measurements carried out on each particle on any direction θ:

$$prob\left(\uparrow_{1q}\right) = \left\langle\Psi\left|\uparrow_{1q}\right\rangle\left\langle\uparrow_{1q}\right|\Psi\right\rangle = \left(1/\sqrt{2}\right)^2 = \frac{1}{2}.$$
$$prob\left(\downarrow_{1q}\right) = prob\left(\uparrow_{2q}\right) = prob\left(\downarrow_{2q}\right) = prob\left(\uparrow_{1q}\right) = \frac{1}{2}.$$

2. Conditional probabilities of outcomes of measurements on either particle, conditional on any particular outcome of any measurement made on the other particle:

$$prob\left(\uparrow_{2q'}/\downarrow_{1q}\right) = \frac{prob\left(\uparrow_{2q'} \wedge \downarrow_{1q}\right)}{prob\left(\uparrow_{1q}\right)} = \frac{1}{2}sen^2\frac{1}{2}qq'.$$

In the specific case θ = θ':

$$prob\left(\uparrow_{2q}/\downarrow_{1q}\right) = 1 = prob\left(\downarrow_{2q}/\uparrow_{1q}\right),$$

$$prob\left(\downarrow_{2q}/\downarrow_{1q}\right) = 0 = prob\left(\uparrow_{2q}/\uparrow_{1q}\right).$$

It is clear then that if we measure both particles' spin along the same direction, the singlet state predicts an anti-correlation between the spin values. If we measure the first particle's spin in the θ direction, and we find the outcome corresponding to "positive" spin ($\uparrow_\theta$), we can predict that the outcome of a later measurement of the second particle's spin in the same direction will be "negative" ($\downarrow_\theta$) with certainty.

In other words, the description offered by the singlet state ψ of the composite system contains the greatest possible amount of information about both systems. By contrast, if we only consider the states of the systems individually taken, W_1 y W_2, we can see that we have lost relevant information. Erwin Schrödinger was perhaps the first to note that "a portion of knowledge of the composite system" is found "squandered on

[4] See Hughes (1989, pp. 149-150), Suárez (2004, appendix 1).

conditional statements that operate between the subsystems". [5] The kind of necessity expressed by these statements, according to quantum mechanics, is nomological merely, and does not seem to be grounded on any physical process. The quantum formalism at least does not describe any physical process capable of transmitting the information required from one system to the other.

A causal explanation of these correlations would have to introduce some type of mechanism, or an additional physical hypothesis to explain these conditional statements. In a causal model the "additional portion of knowledge" would not be "squandered in conditional statements". For instance, in a model where the causes operate directly between the wings of the experiment, the "extra" portion of knowledge could be transmitted directly from one subsystem to the other by means of "mark-transmitters". [6]

2. Correlations, "Forks" and Causal Processes.

In this paper I intend to defend the possibility of a causal explanation of the EPR correlations, one that would satisfy the conditions imposed by e.g. Wesley Salmon's theory of causal explanation. [7] It might help to review some of the main features of this theory. According to Salmon's theory the typical target of a causal explanation (which constitutes its "*explanandum*") is a statistical correlation between two event-types A and B:

$$prob(A \& B) \neq prob(A)prob(B).$$

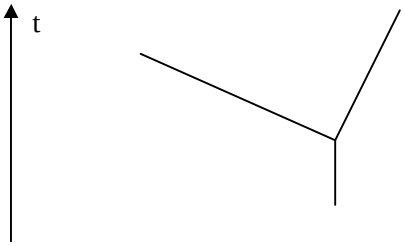
This correlation might be explained by appeal to causal processes and their interactions, which jointly constitute the *explanans*. Salmon defines a *process* as a dynamical object that shows consistency of properties during its existence, and which can be represented as a world-line in four dimensional spacetime. [8] The additional distinction between causal processes and pseudo-processes is notoriously problematic, and has been the object of a large number of subsequent modifications and changes.[9] Inspired by Reichenbach's ideas, Salmon defines a *causal process* as a process capable of transmitting a "mark" – i.e. an alteration in the property of a process that results out of a single interaction. (According to Salmon, an interaction is an intersection of two processes that generates a change in the properties of both).

---

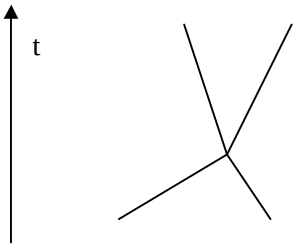
[5] Schrödinger (1933, p. 161).
[6] "Mark-transmitter" is the term employed by both Hans Reichenbach (1956, p. 198) and Wesley Salmon (1984, pp. 148-150).
[7] I take Salmon's 1984 theory because it fits in well with Van Fraassen's conditions. Salmon's theory has since been subject to significant improvements and changes – see Dowe, 2000. But these ulterior developments in general have tended to strengthen logically the conditions on causal explanation. Hence virtually any model that satisfies the conditions of the 1984 theory will ipso facto satisfy those of later theories. This is the pragmatic reason why I presuppose Salmon's 1984 theory throughout this paper – it should not be taken as a defence of this theory against later developments, or other theories of physical causation and explanation.
[8] Salmon (1984, pp. 139-147).
[9] See Salmon (1994).

We then say that a process transmits a "mark" if that change of properties reappears in every point of spacetime contained in the remaining part of the process' world line.

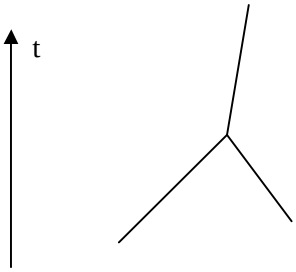
Interactions may be divided in three types, depending on the shape of the intersection that gives rise to them:
Y-interactions:



X-interactions:



λ-interactions:



Interactions are events in spacetime, which have the capacity to effect changes in the properties of certain dynamical objects. We have seen that the object of a scientific explanation, its *explanandum* according to Salmon, is a statistical correlation. Salmon's 1984 theory elaborates further some of Reichenbach's original ideas, and adds the so-called "forks" to the causal processes and causal interactions required for a typical *explanans*.

Conjunctive Fork:

i) P (A&B) > P(A) P(B)
ii) P (A&B/C) = P(A/C) P(B/C)

According to this "fork", in addition to the statistical correlation between events A and B, there is a screening-off relation between the explanatory common cause "C", and the correlated events "A" and "B" that appear in the explanandum. This structure is perhaps more usual in y-interactions, where the type of event at the bifurcation doubles as a common cause. We will see later that this combination (y-interaction with a conjunctive fork structure) is particularly appropriate as a possible application to the EPR correlations.

Interactive Fork:

i) P(A&B) > P(A) P(B)
ii) P(A&B/C) > P(A/C) P(B/C)

In this structure the positive correlation between the events that we are trying to explain does not disappear when conditioned on the common cause. Salmon (1984, pp. 168-174) introduced this type of structure in order to extend his model of causal explanation to cover probabilistic common causes that do not satisfy screening-off.

Perfect Fork:

i) P(A&B) >P(A) P(B)
ii) P(A&B/C) = 1

This structure is simply the limiting case of the two other forks, with a common cause that ascribes probability one to the conjunction of the correlated events- and it is always fulfilled in the case: P(A/C) = 1 = P(B/C).

According to Salmon the explanans in a typical causal explanation has three different components. First, causal processes, which are the means of propagation and transmission of structure and causal order. Second, x-, y- or λ- causal interaction that satisfy the conditions on interactive forks, and help to modify causal structure. And, third, conjunctive common causes, i.e. the specific type of causal interactions that fulfil the conditions for a conjunctive fork, and are mainly responsible for the production of causal structure and order. [10]

Salmon's 1984 theory has proved itself to be a good model of the practice of causal explanation. The objections to the theory – including those powerful objections that have led Salmon himself to abandon the condition of mark transmission in favour of another condition based upon conservation principles – have been conceptual, not empirical. [11] In Salmon's writings there is to my knowledge only one mention of a possible empirical counterexample to the condition of mark-transmission. That is, only one example of a correlation that ought to be causally explained but can't be given such an explanation in accordance to the mark-transmission criterion: the EPR

[10] Salmon (1984, pp. 178-182). Salmon does not make clear whether these three components are necessary for a causal explanation. It seems reasonable to suppose that (the explanans of) any causal explanation contains causal processes and at least one type of fork, conjunctive or interactive.
[11] Salmon (1994).

correlations from quantum mechanics. Salmon adopts Van Fraassen's analysis and writes:

> Is it possible to provide causal explanations of quantum mechanical-phenomena? Van Fraassen argues cogently, on the basis of Bell's inequality and relevant experimental results, that 'there are web-attested phenomena which cannot be embedded in any common-cause model'. [12]
>
> When we ask for the causal (or other sort of) mechanism involved in the production of the EPR correlations, we find ourselves at a loss. [13]
>
> It is this type of remote conservation […] that is perplexing. By what mechanism, we feel compelled to ask, does nature contrive to insure the conservation of momentum in this remote case? [14]

In other words, Salmon believes that the quantum correlations are a worrying case of empirically established correlation that does not admit treatment in accordance to his theory of causal explanation. The reference to Van Fraassen's argument is explicit and heads this worry. I share Salmon's worry: it would not be good news for a theory of explanation by means of physical causes, such as Salmon's, if this theory was not applicable to the most notorious correlations predicted by the empirically most established theory in the history of physics.

One aim of this paper is to show that this suspicion is mistaken, thus rescuing and defending Salmon's theory. Van Fraassen's argument is often interpreted as a refutation of causal realism: this is the metaphysical thesis that states that every statistical correlation between observable or measurable quantities must have a causal explanation. My analysis is intended to show that Van Fraassen's argument can also be employed as a very good overview and guide to the different alternatives and models of causal explanation. Specifically I aim to show that it is possible to offer models or causal explanations for the EPR correlations. Salmon himself wisely left an open door to this possibility:

> It would be premature, I believe, to conclude that causal explanations of quantum phenomena are impossible in principle […] The nature and role of causality in microphysics is a deep and difficult matter to sort out. [15]

3. Causal Models for EPR Correlations

Let us now return to our initial theme – the EPR correlations – as graphically represented in figure 1. It is possible in principle to provide two types of causal explanations for a correlation between two event-types A and B: a direct causal relation between A and B, and a common cause structure, with a common cause C that underlies the correlation.

[12] Salmon (1984, p. 254).
[13] Salmon, (ibid, p. 251).
[14] Salmon (ibid, p. 256).
[15] Salmon (ibid, pp. 254-5).

3.1. Direct-Cause Models

These models assume that the outcome-event in one wing of the experiment (s_1) is a direct cause of the outcome-event in the other wing (s_2):

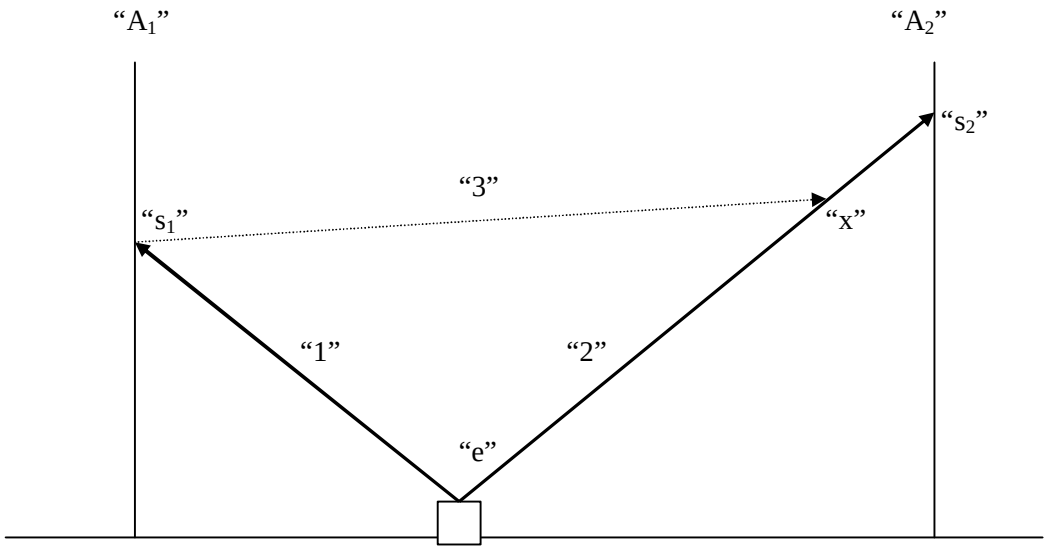


Figure 2

In this spacetime graph A_1 and A_2 represent the worldlines of the measurement devices; "1" y "2" represent those of the particles; and the line comprised between s_1 and x is the worldline of a direct causal process between the wings of the experiment. s_1, s_2, x and e denote event-types, where e is the particles' emission event, and s_1 y s_2 are the outcome-events that result from measurements on particle "1" by device A_1; and on particle "2" by device A_2, respectively. x is the reception event by particle "2" of the causal influence emitted by particle "1"; and it constitutes a partial cause of the outcome s_2.

This model satisfies the conditions on causal explanation imposed by Salmon's 1984 theory. The explanation of the correlation between s_1 y s_2 is given by four causal interactions and five partial causes (event-types that constitute a necessary contribution to the explanation of the correlation):

"e": y-interaction which produces both particles in state ψ.

"s_1": λ-interaction between device "A_1" and particle "1" which results in: i) outcome-event "1", and emission of causal influence "3".

"x": λ-interaction between the causal influence "3" and the particle "2" which changes the state of the particle.

"s_2": λ-interaction between the particle "2" and the measurement device "A_2", resulting in outcome-event s_2.

"e": partial common cause of s_1 and of x.

"A_1": partial cause of s_1. "A_2": partial cause of s_2.

"s_1": partial cause of x. "x": partial cause of s_2.

3.2. Common-Cause Models

There is yet another type of causal model, which asumes that the causal explanation of a correlation between events s_1 and s_2 is not the outcome of a direct causal influence but is given by a common cause "c". This is connected by means of a series of causal processes with both events, and is implied to satisfy the following requirements:

i) The common cause c is the emission event of both particles at the source.
ii) The common cause lies in the past lightcone of both events s_1 and s_2.
iii) The events c, s_1 and s_2 constitute a conjunctive fork – since c screens off s_1 from s_2: $P(s_1 \& s_2 / c) = P(s_1 / c)\, P(s_2 / c)$.
iv) The causal influence of c on s_1 y s_2 is transmitted by the particles themselves along their trajectory.

This common cause model can be graphically represented thus:

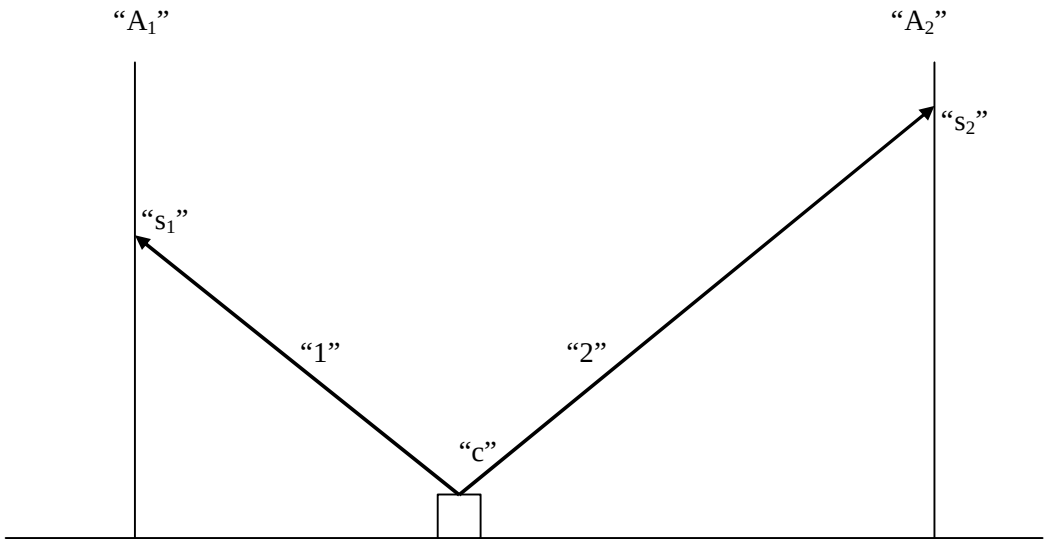


Figure 3

In this diagram A_1 and A_2 again represent the measurement devices' worldlines; "1" and "2" those of the particles, as well as the causal processes that transmit the corresponding mark. s_1, s_2 and c denote event-types, where "c" is the emission event and the partial common cause of s_1 and s_2, the outcome events on each wing.

This model too obeys the requirements laid out by Salmon in 1984. The explanation of the correlation between s_1 and s_2 is given by three causal interactions and three partial causes:

"c": y-interaction that generates at *c* two particles in state ψ.

"s_1": λ-interaction between the measurement device "A_1" and the particle "1" that results in the outcome-event s_1.

"s_2": λ-interaction between the measurement device "A_2" and the particle "2" that results in the outcome-event s_2.

c: type-event that constitutes the partial common cause of both s_1 and s_2,

A_1: partial cause of s_1.

A_2: partial cause of s_2.

4. Arguments Against Causal Models for EPR

There are as a matter of principle two different types of causal models that we might construct for the EPR correlations, both *prima facie* agreeing with the requirements imposed by Salmon. However, as already mentioned, Salmon expresses a strongly felt pessimism for the possibility of any causal mechanism for EPR. What exactly underlies this pessimism? In this section I would like to review briefly some of the main features of the most important arguments that have been presented against these models of EPR, namely Bas Van Fraassen's; these arguments are often quoted by Salmon himself. In a later section I will reply to these arguments, providing a number of alternatives in order to defend the applicability of Salmon's theory of causal explanation to the EPR correlations.

4.1. The Relativistic Objection Against the Direct-Cause Model

The most important objection against the direct-cause model is related to the theory of special relativity. Salmon leans heavily on Van Frassen's argument against what Van Fraassen calls an "*explanation by coordination*" of the EPR correlations. Van Fraassen's argument is a brief allusion to the superluminal character of the speed of the causal influence (represented by the worldline "3" in figure 2): [16]

> By *coordination* I mean a correspondence effected by signals (in a wide sense): some energy or matter travelling from one location to another, and acting as a partial producing factor for the corresponding event. The situation need not be deterministic – there can be indeterministic signalling if the signal is not certain to arrive and / or not certain to have the required effect. But the word "travel" must be taken seriously. Hence this explanation cannot work for corresponding events with spacelike separation. To speak of instantaneous

[16] Van Fraassen (1982, in the expanded 1989 version, p. 112)

travel from X to Y is a mixed or incoherente metaphor, for the entity in question is implied to be simultaneously at X and at Y – in which case there is no need for travel, as it is at its destination already.

Let us analise this line of argument carefully. As we have already seen, figure 2 represents a causal connection between the two wings of an EPR experiment. But we have already seen that in a typical EPR experiment, the outcome-events s_1 and s_2 are spacelike connected; so the causal influence "3" must travel at speed greater than light for at least a part of its trajectory. Now, let us suppose that the trajectory is exactly as in figure 2, with the reception of the influence at x some time after its emission at s_1, in the frame of reference with respect to which the laboratory is at rest.

According to the special theory of relativity, any inertial frame in movement with respect to the laboratory must be equally valid for describing the causal mechanism that explains the EPR correlations between s_1 and s_2. It is a well known fact that, according to the theory, the temporal order of two spacelike connected events is neither absolute nor independent of the frame of reference. There will therefore exist inertial frames of reference that will describe the mechanism very differently. For instance, in some frames the time lapsed between events s_1 and x is longer than that which lapses between those same events in the laboratory frame of reference:

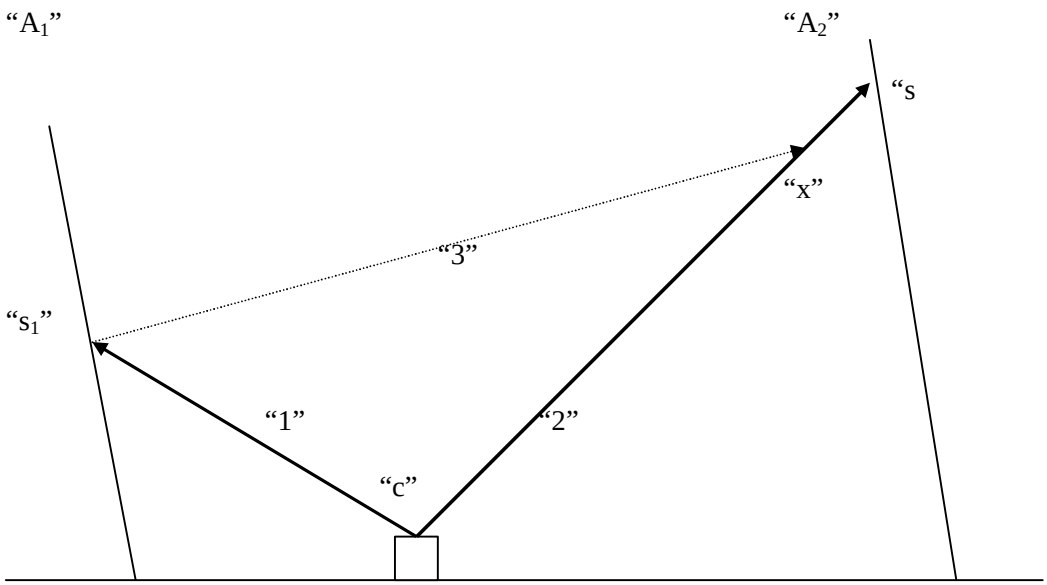


Figure 4.

However, there are other inertial frames of reference, equally valid for the description of physical events, in which event s_1 will be seen to occur *after* event x:

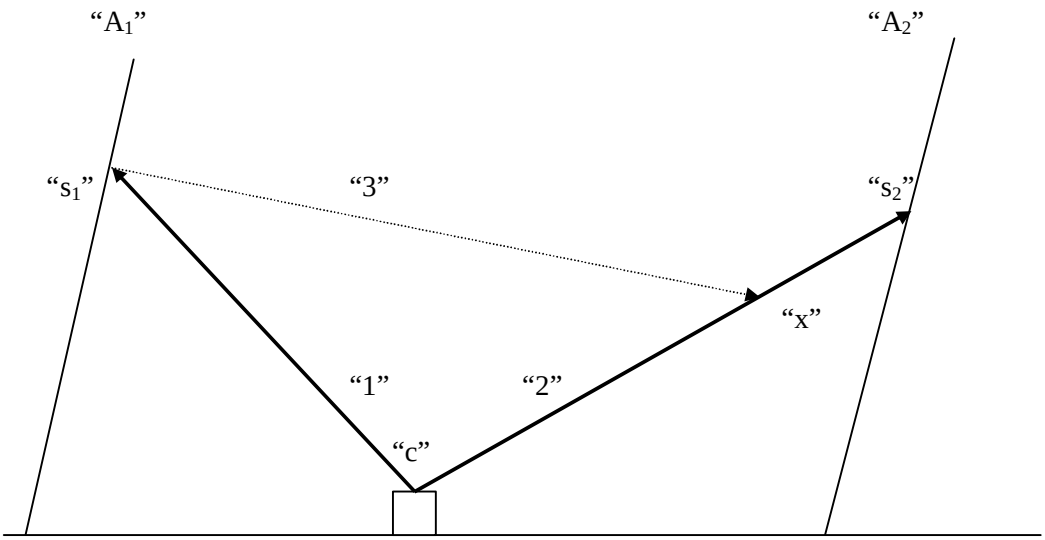


Figure 5.

The existence of two inertial frames of reference that invert the temporal order of two events also entails the existence of a tirad frame of reference – intermediate between the two – in which the reception of the causal influence is instantaneous with its emission:

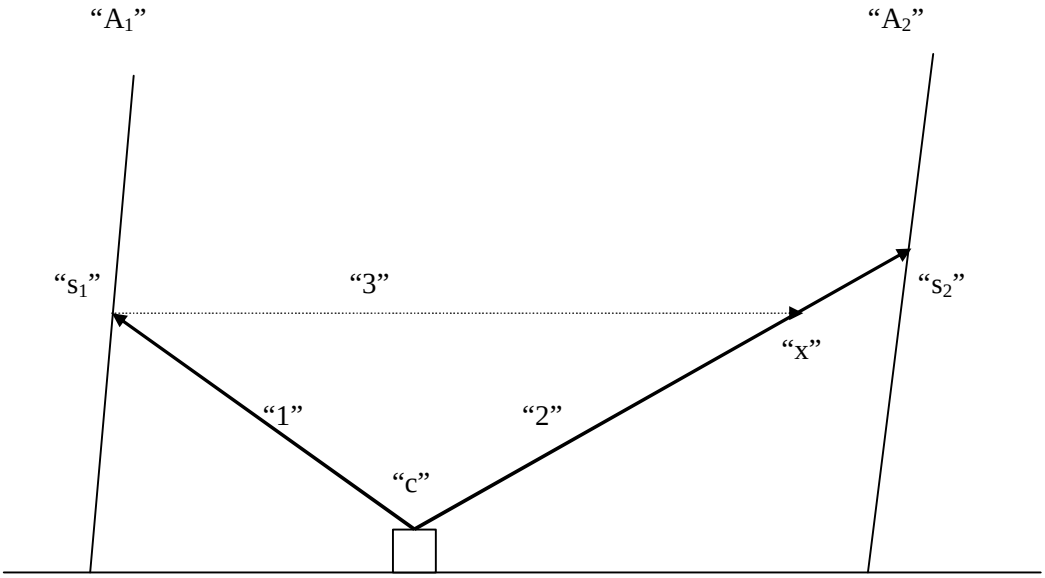


Figure 6.

This is the scenario that worries Van Fraassen, for it implies the instantaneous coexistence of distant cause and effect, which appears to be incoherent with one seemingly necessary property of a causal influence, namely that it "travels".

4.2. Van Fraassen's argument against Common Cause Models

Van Fraassen's argument against direct-cause models is thus very brief. The main part of his 1982 article is devoted to refuting the possibility of a common cause model, of the form shown in figure 3. Van Fraassen's argument against common causes is based upon Reichenbach's so-called *principle of common cause*: "If an improbable coincidence has occurred, there must exist a common cause". [17]

Two comments are in order regarding this principle. First, by "improbable" Richenbach does not simply mean a coincidence with a low prior probability. Rather, what he has in mind is a coincidence between two event-types *A* and *B*, which we do not have any reason to suppose have a direct mutual causal dependence, that is we have no reason to suppose that *A* causes *B* or *B* causes *A*. Secondly, according to Reichenbach, it is not the case that a coincidence is genuine if and only if it is lawlike, in some metaphysical sense. On the contrary, for Reichenbach a coincidence is genuine if it is an example of an empirically established correlation between two event types.

The EPR correlations satisfy such requirements fully since they are experimentally verified and, in addition, are a precise prediction of quantum mechanics. The improbable coincidence is nothing but a statistical correlation between the outcome-events of spin measurements in each wing of the experiment:

$$prob(s_1 \,\&\, s_2) > prob(s_1)prob(s_2).$$

Van Fraassen's argument introduces an additional complexity, since it considers the possible outcomes of spin measurements made on the particles with measuring devices' settings in different directions. The experimenter must choose the direction of spin measurements on both particles, which need not agree in both wings. Let us suppose that the direction of measurements on the first particle is given by a certain angle θ, and the direction of measurement on the second particle by θ'. Van Fraassen denotes the setting-event by the first experimenter as "*a*" and that in the second experimenter as "*b*". It should be clear that, in all frames of reference, *a* precedes s_1, and *b* precedes s_2.

Van Fraassen's argument aims to show that the Bell inequalities, and their experimental violation, imply that a common cause model for the EPR correlations is impossible. Bell employs a notorious probabilitistic condition, *factorizability*, as a condition of physical locality:[18]

[17] Reichenbach, 1956, p. 157 ff.

[18] Bell (1964, pp. 195-6), reprinted in Wheeler and Zurek (1983, pp. 403-4.). Bell's identification of factorizability with physical locality has been the object of a fascinating debate and controversy – see e.g. Suárez (2000) for a review and some references. Van Fraassen's terminology should be equally controversial, since his terms ("causality", "autonomy", "locality") suggest physical or causal facts, when the truth conditions of these three expressions are strictly probabilistic.

"Factorizability":

$$prob(s_1 \,\&\, s_2 / a \,\&\, b \,\&\, \Psi) = prob(s_1 / a \,\&\, \Psi)\, prob(s_2 / b \,\&\, \Psi).$$

Van Fraassen analizes this condition as a conjunction of three logically independent further conditions, which he calls "causality", "hidden locality", and "hidden autonomy", and which he defines as follows:

"Causality":

$$prob(s_1 / s_2 \,\&\, a \,\&\, b \,\&\, \Psi) = prob(s_1 / a \,\&\, b \,\&\, \Psi),$$
$$prob(s_2 / s_1 \,\&\, a \,\&\, b \,\&\, \Psi) = prob(s_2 / a \,\&\, b \,\&\, \Psi).$$

This condition is a direct application of Reichenbach's screening off. It states that the conjoint event ($a \,\&\, b \,\&\, y$) makes event s_2 statistically irrelevant to the probability of s_1, and viceversa. It has already been noted that, according to Reichenbach and Salmon, screening off is a necessary condition on a conjunctive fork. According to this point of view, if this condition is not satisfied, then the conjunction ($a \,\&\, b \,\&\, y$) can not be a common cause of the conjunctive fork variety.

"Hidden Locality":

$$prob(s_1 / a \,\&\, b \,\&\, \Psi) = prob(s_1 / a \,\&\, \Psi),$$
$$prob(s_2 / a \,\&\, b \,\&\, \Psi) = prob(s_2 / b \,\&\, \Psi).$$

This condition is also a screening off condition, which now states that the conjunction of the state ψ of both particle, with the setting-event in the direction of measurement in the nearby wing makes the setting-event in the far away device irrelevant to the probability of the outcome-event in the nearby wing.

"Hidden Autonomy":

$$prob(\Psi / a \,\&\, b) = prob(\Psi).$$

This condition establishes that the probability of the particles to be in a particular state ψ at the time of their emission is independent of the selection of the setting-events in either wing.

Van Fraassen then shows that the conjunction of *causality*, *hidden locality* and *hidden autonomy* implies Bell's factorizability condition, which in turn gives rise to the inequalities violated by experimental results as well as the quantum mechanical predictions. Since factorizability must be false, one of these three conditions (or some logical conjunction of them) must similarly be false.

Van Fraassen provides us with the following argument in favour of *hidden autonomy* and *hidden locality* – with the ultimate end in mind to place the blame on *causality*: [19]

[19] Van Fraassen (1982, in its corrected and extended version 1989, p. 105).

> If the probability of a given outcome at [one wing] is dependent not merely on the putative common cause, but also on what happens at [the other wing], or if the character of that putative common cause itself depends on which experimental arrangement is chosen (even after the source has been constructed) then I say that the two-outcome events have not been traced back to a common cause which explains their correlation.

In other words, according to Van Fraassen there might be common-cause models in which *hidden locality* and / or *hidden autonomy* fail, but these models would not constitute common cause *explanations* of the correlations. If on the other hand *causality* fails then common cause models are just not viable, since this condition embodies the application of Reichenbach's principle. Van Fraassen then goes to argue that the EPR correlations themselves show *causality* to be false. This is because the state of the particles at the time of their emission does not screen off the outcome-events from each other. For, let us suppose that $a = b = \theta$, without loss of generality. Then *causality* reduces to:

$$prob(s_1/s_2 \,\&\, \Psi) = prob(s_1/\Psi),$$
$$prob(s_2/s_1 \,\&\, \Psi) = prob(s_2/\Psi).$$

And this condition is certainly false, since according to quantum mechanics:

$$prob(s_1/s_2 \,\&\, \Psi) = 1 \neq prob(s_1/\Psi) = 1/2,$$
$$prob(s_2/s_1 \,\&\, \Psi) = 1 \neq prob(s_2/\Psi) = 1/2.$$

Van Fraassen thus concludes that a common cause model for EPR of the sort envisioned by Reichenbach is not viable: [20]

> "The conclusion is surely inevitable: there are well attested pehnomena which cannot be embedded in any common-cause model".

5. Replies and Alternatives

My exposition so far is already geared towards a particular conclusion. In particular in the last section, in reviewing the arguments against causal models for the EPR correlations, I have been anticipating several replies and possible alternatives. In this section I point out these alternatives explicitly and develop them into sketches of a range of possible causal models of EPR. I leave for later work the technical development of these sketches into mathematically sophisticated causal models. In this paper I only aim to argue that all these models (of either the direct or common cause varieties) have not yet been shown inconsistent with the experimental results. Of course this does not amount to a proof that one among these models is a literal and true description of the causal mechanism underlying the EPR correlations; it only

[20] Van Fraassen (1982, in the 1989 version, p. 108).

entails that there is no proof to the contrary. This conclusion should not be surprising, in my view: causal models are generally metaphysical or ontological interpretations of physical phenomena, and it ought not surprise us to find that such interpretations are under-determined by the empirical outcomes, i.e. by the statistical correlations between observable or measurable quantities.

5.1. In Favour of Direct-CauseModels

There are at least three different replies to the relativistic argument against direct-cause models. The first considers the possibility of causal influences without a material or physical basis; the second questions the requirement that quantum phenomena be Lorentz-invariant; the third admits the possibility of backwards causation in time. Let us consider them in turn.

5.1.1. The causal influence might not have a material basis

It is true that there exists a frame of reference where a direct-causal influence between the wings will be instantaneous (see figure 6). However, an instantaneous causal influcence need not be incoherent. Van Fraassen is right to claim that if a physical object is in two locations in space at the same time, it would not make sense to suppose that it has "travelled" from one location to the other. But the direct cause model does not necessarily imply that there is a displacement of a *physical* or *material* object from one wing to the other. The *causal influence* is transmitted from one location to the other: this only implies that a material object is in two places at once if the causal influence can only be transmitted by means of the displacement of a physical object.

There are several sophisticated and well known theories of causation available nowadays, which do not require the existence of a physical object in order to transmit the causal influence.[21] As a matter of fact in many of these theories, the "instantaneous transmission of causal influence", suitably understood, is not only possible, but desirable (suitably understood – the terms "transmission" or "influence" should be understood as devoid of a physical connotation). In recent years, several convincing arguments have been developed in favour of the claim, specific to quantum mechanics, that the causal influence that might be operating in EPR does not transmit itself through any material displacement. [22]

To sum up, the first reply is that causal influence need not require material transmission or physical displacement. If so, Van Fraassen's argument, based upon the impossibility of an instantaneous physical displacement would have no consequences. (This reply also allows *prima facie* for the compatibility of a direct-cause model with special relativity). However, it must be admitted that on this interpretation the direct-cause model would not be a genuine causal explanation according to Salmon's theory. This is not a counterfactual theory of causation, but a physical theory, which seeks a material or physical entity – a causal process – as the basis of the transmission of any causal influence.

---

[21] Perhaps the best known is the counterfactual theory due to David Lewis (1986). Psillos (2002, cap. 3) is a clear and concise introduction to this type of theory.
[22] Maudlin (1995, cap. 5) is an excellent defence of this point of view.

5.1.2. Causal processes might be non-relativistic

The second reply denies that the causal model that explains the EPR correlations need be Lorentz-invariant. Why should we suppose that quantum processes must obey the special theory of relativity? Several extant interpretations of quantum mechanics already clearly abandon the requirement of relativistic invariance at the ontological or causal level. A very well known case is Bohm's theory, whose wave equations (corresponding to the "quantum field") are not invariant under the Lorentz transformations.[23] Another widely debated and established case is the modal interpretation, which postulates stochastic transitions between the so-called value states, which are not Lorentz-invariant.[24]

If we abandon the commitment to Lorentz-invariance, we could go on to suppose that the causal process that connects both wings of the experiment determines a privileged frame of reference; any other frame would be a fictitious one. In Jim Cushing's (1995) terminology, the theory of relativity would be a phenomenological theory, not a fundamental one, with respect to quantum causal processes. The statistical frequencies of the EPR experiments, including the correlations, would be invariant under the Lorentz-transformations, but the causal processes that explain these correlations would not be so. The determination of the privileged frame of reference might become an empirical matter. We could provisionally suppose that the privileged frame is the one in which the first measurement device is at rest, corresponding to figure 2. There does not seem to be a reason, from the perspective of Salmon's theory of causal explanation, to deny this second possibility.

5.1.3. Causal influences might be transmitted backwards in time

The third reply would insist on the Lorentz-invariance of quantum causal processes, and would accept fully its consequences. The main consequences, as we have already seen, are the existence of a frame in which the causal influence is instantaneous (figure 6), and other frames in which the influence travels back in time (figure 5).

Let us consider both consequences in turn. The first entails the incoherence noted by Van Fraassen, which we can only get around now by abandoning the familiar meanings associated with "transmission", "travel", and "influence". The special theory of relativity would imply that the "transmission" of causal influence is not an objective physical fact of the situation, since whether there is "transmission" in this sense or not is a frame-dependent issue. But this does not imply that the existence of the causal process between the outcome-events is frame-dependent, on Salmon's definitions. For it was already noted (in section 2) that these terms have a rather technical meaning in Salmon's theory. It is still the case that in every frame – including the frame with instantaneous causation – there is a four dimensional worldline that exhibits consistency of properties, and which transmits a mark in the technical sense invoked by Salmon, and explained in section 2 above. We would still have a perfectly legitimate causal process in that frame – what would be illegitimate is

[23] Cushing (1995) also emphasises the distinction between fundamental and phenomenological invariance. According to Cushing, Bohm's theory is committed to a phenomenological reading of the special theory of relativity which does not require fundamental invariance. This is a reading congenial to my second alternative proposal for interpreting the direct-cause model for EPR.
[24] See Dickson and Clifton (1998).

to describe this process as a "transmission" of "influence" in the familiar sense of these words.

The second consequence that this third line of reply would have to accept is the fact that the causal influence would travel back in time in some further frames of reference. The argument in the previous paragraph comes to show that there is no reason why we should not speak of "causal processes transmitting marks" in these frames, as long as we stick to Salmon's technical definitions. But there is certainly an added difficulty which relates to the coherence of backwards in time causation. Does the concept make sense? Can it apply to EPR? Fortunately these questions have positive answers. For a start the traditional arguments against backwards in time causation have been long refuted.[25] In addition, several causal models for the EPR correlations that make use of the hypothesis of backwards in time causation have already been developed. These models go further than the present alternative since they make use of backwards in time causation *in the most familiar* reference frame: the frame with respect to which the laboratory itself and all measurement devices are at rest.[26]

It is of course possible that these models, which use the hypothesis of backwards in time causation, will eventually be proved implausible or extravagant; but an empirical *refutation* of these models does not seem likely. (In my opinion, these models are as plausible as any other model or interpretation of quantum mechanics presently debated; and they would in any case be at least as plausible as the hypothesis according to which the correlations are a brute fact of nature, lacking in any explanation). Contrary to what Salmon and Van Fraassen claim, it is still perfectly possible to explain the correlations by means of direct-causal processes that in every way agree perfectly with the requirements of Salmon's theory.

5.2. In Favour of Common Cause Models

Let us quickly review Van Fraassen's argument against common cause models. Its first premise is (Pr1): *hidden locality* and *hidden autonomy* could hold in a common cause model, but these models would not explain the correlations. The other premises are (Pr2): *causality* must necessarily hold for such a model to be viable, and (Pr3): The violation of Bell's inequalities itself tells us that causality is actually false. The conclusion Van Fraassen draws is thus (iv): no common cause model for the EPR correlations is viable. I will argue that the argument is valid but not necessarily sound: the premisses are all controversial and (Pr1), in particular, seems likely to be false.

We have seen how the common cause models that Van Fraassen considers and rejects make several assumptions (assumptions (i-iv) in section 3.2.) Van Fraassen does not make explicit at least three of these assumptions (assumption (iii) is made explicit to the extent that it is required that any common cause must be a screener off). However, neither of these assumptions are equivalent to – nor a logical consequence of – any of

[25] By Dummett (1954) and Horwich (1987) among others.
[26] Costa de Beauregard (1977) is one of the first EPR models of this kind. Currently the best-known model of this sort is probably Huw Price's (1996, cap. 9). In this model the setting-event in each wing of the experiment is a cause of the state of the particles as they are emitted. (As I point out later on, this model violates Van Fraassen's *hidden autonomy*).

the requirements established by Salmon's theory for a causal explanation. The following question arises: Would it be possible perhaps to construct causal models for EPR that satisfy Salmon's requirements but do not satisfy any of these additional assumptions? Let us carefully scrutinise the different possibilities.

5.2.1. The common cause might be an interactive fork

Let us begin by considering assumption (iii) which is arguably the most controversial, in addition to being the only one explicitly acknowledged by Van Fraassen. Van Fraassen insists, following Reichenbach, that any common cause C of two correlated effects A and B must obey the *screening off* condition:

$$P(A\&B/C) = P(A/C)\,P(B/C)$$

According to both Reichenbach and Van Fraassen, *screening off* is a necessary condition – albeit not sufficient – for a common cause. Why?. We have already seen that Salmon himself allows for interactive as well as conjunctive forks. Can a common cause fail to satisfy screening off? Some philosophers think so in the case of genuinely probabilistic causes – for example Nancy Cartwright.[27] If Cartwright is right then premise (Pr2) in Van Fraassen's argument is false. I will not take sides on this issue – which is at the heart of the contemporary debate related to the Markov Causal Condition, a debate that does not fall within the scope of this paper.[28] But it is worth pointing out that if we abandon Reichenbach's screening off requirement then the common cause of the EPR correlations could be the very state of the particle pair at the time of their emission considered and rejected by Van Fraassen. [29]

5.2.2. The common cause might not be the emission event

Assumption (i) could also be false – and Van Fraassen's argument could consequently be invalid. This would happen for instance if the common cause was an event prior to the emission event of the particles (for instance the creation event of the source itself, or the preparation event of the particle pair at the source). As far as I know this type of model has never been studied or proposed. But it is hard to understand why, since it constitutes a typical hidden variable model. What's more – it could be constructed in such a way as to satisfy the screening off condition, like a classical conjunctive fork.

Perhaps the reason why this model has never been developed is the inherent difficulty in identifying some event *d* in the proper past of *c*, s_1 y s_2 that can play the role of a hidden common cause. But perhaps the reason has rather to do with the suspicion that if this event or hidden variable was to satisfy Reichenbach's conditions, it would necessarily have to satisfy the factorizability condition, which would imply a commitment to the Bell inequalities. This reasoning does not strike me as convincing, however, for the following reason. There does not seem to be any reason why, if there was a common cause or hidden variable, this cause could not be a partial cause of each of the outcome-events, and *in addition* a partial cause of the setting-events *a* and *b* (although these events lie outside the future light cone of *c*, they can nevertheless lie

[27] Cartwright (1988).
[28] See, for instance, Hausman and Woodward (1999). Some of the issues debated have already had application to EPR – by, for instance, Hofer-Szabo, Redei and Szabo (1999).
[29] Cartwright (1989, cap. 9) sketches a model of this type.

within the future light cone of *d*, since *d* is in the proper past of *c*). Moreover, *d* could be a partial cause of the emission event *c*, which could in turn be a partial cause of the outcome-events. In other words, there does not seem to be any reason to exclude a complex causal structure of the following type:

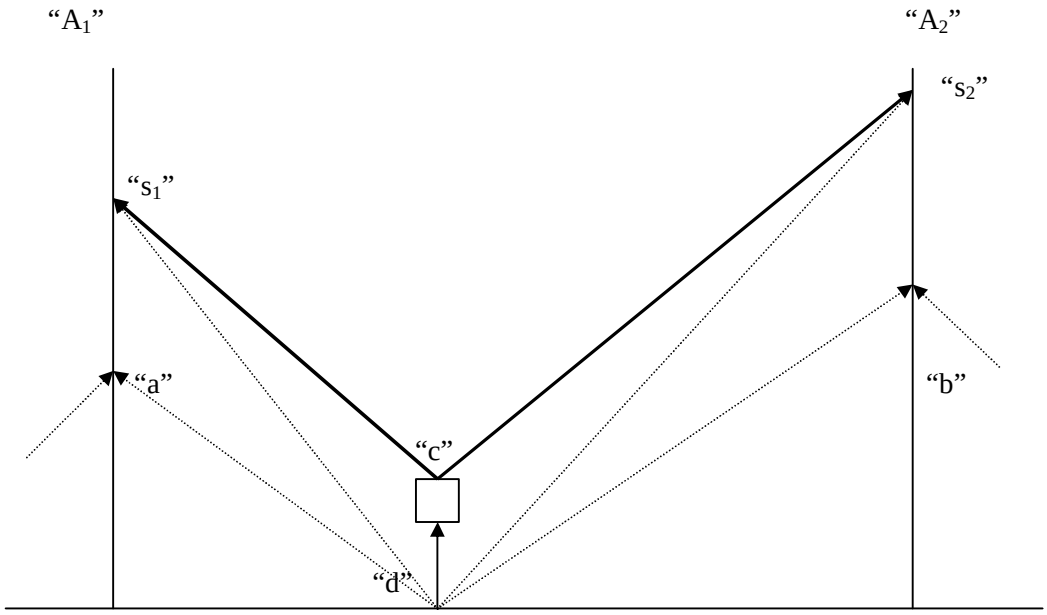


Figure 7

In this causal structure: *d* is the partial common cause of *c*, *a*, *b*, s_1, s_2; *c* is the partial common cause of s_1, s_2; *a* is the partial cause of s_1; *b* is the partial cause of s_2. These "common causes" can not be expected to form a conjunctive fork by themselves. That is:

$$prob(s_1 \,\&\, s_2 / c) \neq prob(s_1 / c)prob(s_2 / c),$$
$$prob(s_1 \,\&\, s_2 / d) \neq prob(s_1 / d)prob(s_2 / d).$$

However, the conjunction of both might (or might not) satisfy factorizability:

$$prob(s_1 \,\&\, s_2 / a \,\&\, b \,\&\, c \,\&\, d) = prob(s_1 / a \,\&\, c \,\&\, d)prob(s_2 / b \,\&\, c \,\&\, d).$$

In this case, Bell's inequalities are not derivable with respect to each isolated common cause, even if they might be derivable with respect to the logical conjunction of both.[30]

5.2.3. The common cause might lie outside the light-cone

[30] In other words: whether or not this model is empirically adequate is a distinct and different question to the one addressed by Van Fraassen's argument regarding the empirical adequacy of a model that has only *c* as a possible common cause.

The second supposition might also be false. The common cause might not be a discrete event, and need not lie outside the light cone of s_1 and s_2. Specifically, the common cause might be the state of the two particles just prior, in the laboratory frame, to the first outcome-event. Since the two particles are physically separated from each other, the spacetime location of the state must be a hyperplane, or at least a (presumably closed) region of such a hypersurface that contains the physical position of both particles at that time just prior to the first outcome-event.[31] For instance:

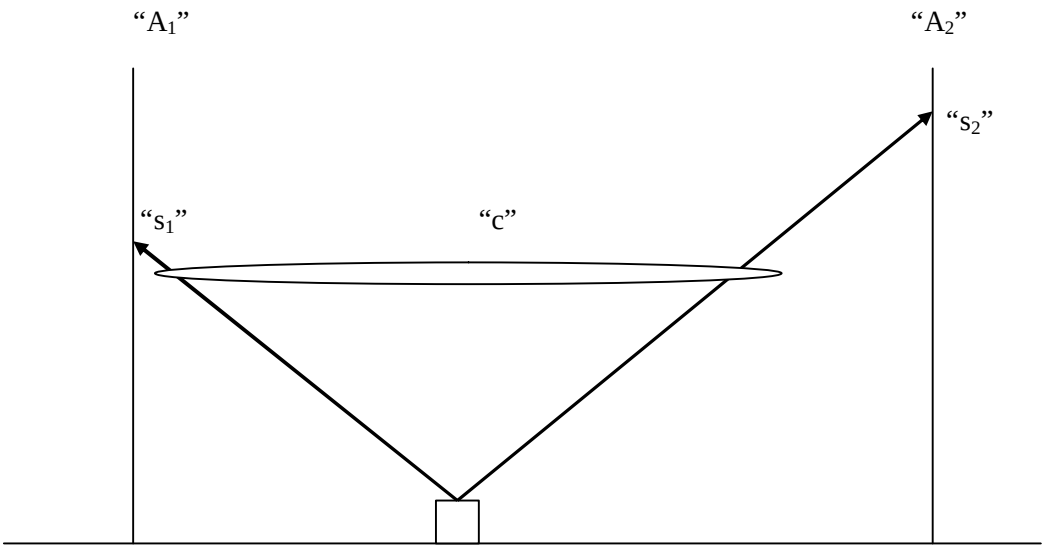


Figure 8

5.2.4. The causal influence might not be transmitted with the particles

The causal influence's trajectory might not correspond to that of any of the particles. For instance, it is possible to suppose that, even if both particles' trajectories were classical and continuous, the causal influence could exhibit important discontinuities in space and / or time. It is not necessary here to broach the details of this option. [32] It is enough to mention that such a supposition is not alien to the history of quantum mechanics. On the contrary – it seems to agree with the spirit of the orthodox interpretation of quantum mechanics, the Copenhagen interpretation, according to which it is not legitimate to ascribe a classical or continuous trajectory in spacetime even to quantum particles themselves.

5.2.5. "Hidden autonomy" might be false

I have already mentioned (footnote on page 25) that this condition is false in the models of backwards in time causation of Costa de Beauregard and Huw Price. In these models (sometimes referred to as "zigzag models") the setting-event in any of the wings is causally related to the state of the particles as they are emitted at the

[31] Jeremy Butterfield (1989) explores some of these possibilities.
[32] I discuss it in Suárez (2000c, p. 9). Chang and Cartwright (1993) develop a model of this kind.

source. Specifically the setting-events are a partial cause of the initial state of the particles – an earlier event in the laboratory frame:

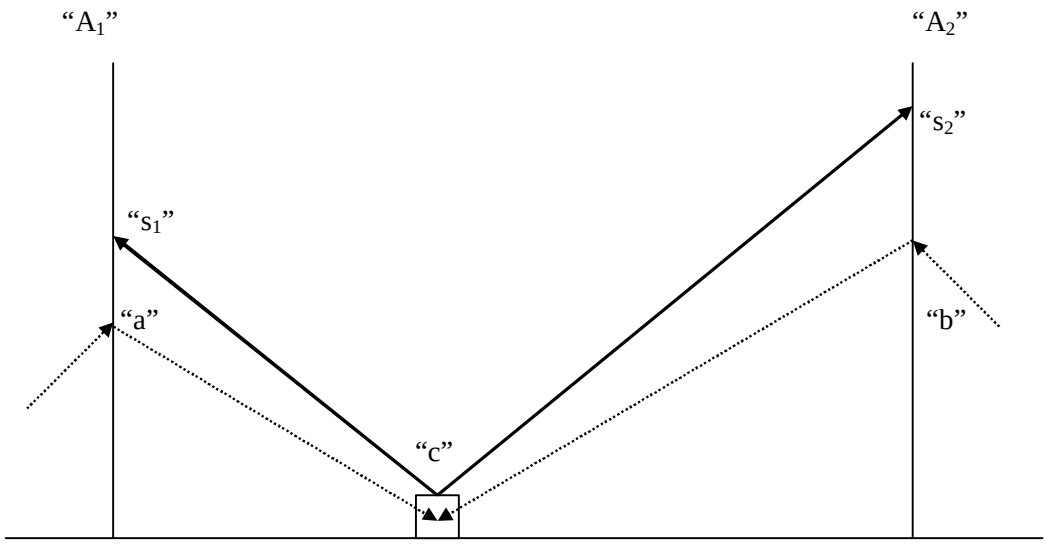


Figure 9

In other words, *hidden autonomy* is false in these models:

$$prob(\Psi/a\,\&\,b) \neq prob(\Psi).$$

Van Fraassen's premise (Pr1) entails that these models can not provide explanations of the correlations. The plausibility of these models can indeed be questioned; but in so far as they are plausible, they seem *prima facie* as *explanatory* of the correlations as any other common cause model.

5.2.6. "Hidden locality" might be false

There is a further causal possibility. Van Fraassen's *hidden locality* can also be false in a legitimate common cause model of the EPR correlations. It is possible to suppose that a setting-event in one wing is a partial cause of the outcome-event in the opposite wing. In such a case:

$$prob(s_1/a\,\&\,b\,\&\,\Psi) \neq prob(s_1/a\,\&\,\Psi),$$
$$prob(s_2/a\,\&\,b\,\&\,\Psi) \neq prob(s_2/b\,\&\,\Psi).$$

As I already noted Van Fraassen claims that if this condition were false then "the two-outcome events have not been traced back to a common cause which explains their correlation".[33] Van Fraassen is quite right that, in a model that violates *hidden*

[33] Van Fraassen (1982, in the 1989 version, p. 105).

locality, the correlation can not be explained *only* by means of a common cause. But it does not follow from that there can not exist a common cause as a part of the full causal structure that explains the correlations. For instance, the outcome events might be partially caused by the action of the common cause, and partially caused by the setting-events. In other words, Van Fraassen's reasoning is not meant to exclude the following causal structure, which is perfectly compatible with Salmon's theory of causal explanation:

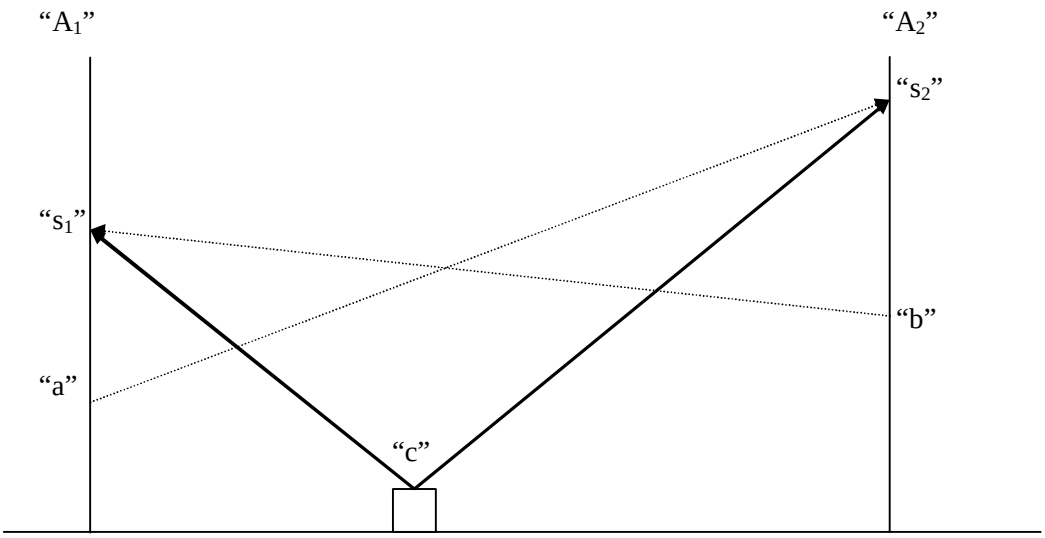


Figure 10

5.3.The Status of Causal Realism

In spite of efforts to carefully exclude all causal explanations of the EPR correlations, the alternatives available – in the form of several models – are many and diverse. Some of these causal models are more plausible than others; but none has so far been excluded by the experimental results. Experiments might be designed in the future to refute empirically one or another particular model. For that task Van Fraassen's argument is an excellent tool, since it constitutes a useful guide for sorting out the models into different classes. It seems very implausible though that all causal models of the EPR correlations will ever be refuted. It would arrogant on my part to claim to have given a full account of the possibilities, and there are no limits to the imagination in this regard. For instance, we could go on to consider all kinds of hybrid proposals: direct causal relations between the wings superposed with common causes, etc. All of these alternatives seem compatible with Salmon's theory of causal explanation, since they explain the correlations on the basis of a combination of causal processes, different types of causal interactions, and forks. Van Fraassen's argument seems to have induced on Salmon an unfounded pessimism regarding the prospects of the latter's theory of causal explanation.

Perhaps one day the thesis of causal realism will be shown to be untenable: There are some interesting and persuasive philosophical arguments to the effect that causal realism is neither necessary nor inevitable.[34] But "dispensible" and "false" are different properties. It seems to me to be part of the metaphysical character of a metaphysical thesis that it can not be empirically refuted. So it does not seem unreasonable to suppose that no experimental reasons could ever force us to abandon causal realism.[35]

6. Propensities and Causal Models

In this final section I would like to broach briefly a further issue, in order to display one way in which Van Fraassen's argument can be a very useful tool. My claim – which there is only space to develop in its bare essentials here – is that the classification of causal models that has emerged in the previous section can help distinguish different notions of quantum propensities and their features. Let us assume that whatever causal model underlies EPR is in effect a description of the workings of quantum propensities; i.e. let us assume that the relata of these causal relations are (often, not always) propensities. (This is of course an additional assumption to Wesley Salmon's theory – which does not *per se* require propensities to constitute the causal relata of any causal laws or relations – but as I intend to show in what follows, it is compatible with it). Thus the probability distributions that appear in these models are nothing but the manifestations or displays of quantum dispositional properties, or propensities. Specifically: the probability distributions over the outcomes of spin measurements on each wing are the manifestations or displays of quantum propensities in particular experimental set-ups.[36] What would the causal models studied in section 5 imply regarding these dispositions or propensities?

At this point I need to make explicit one assumption regarding the underlying ontology of dispositional properties, or propensities, that I will be employing. The assumption is neither necessary nor universal, but it is in line with the bulk of the philosophical literature on dispositions: [37] Dispositions, and propensities, are not properties of the experimental set-ups designed to test them, but rather properties of the systems themselves under test. The ascription of these properties is thus independent of whatever experimental arrangement these systems find themselves in. These arrangements simply serve to extract manifestations, or displays, of dispositions or propensities, but do not determine their possession by any entity or system. Hence each individual EPR particle will possess propensities of its own

[34] Fine (1989) is an excellent defence of the dispensability of causal explanations for EPR, without attempting an empirical refutation of such explanations.
[35] It is in my view puzzling that Van Fraassen – whose constructive empiricism is so openly sceptical regarding the possibilities of any metaphysics – would invest such effort in the empirical refutation of a metaphysical thesis. It seems to me that the natural position for a constructive empiricist to adopt is the one I adopt in this paper: a combination of suspension of judgement (analogous to religious agnosticism) regarding the ultimate structure of reality; and open scepticism about our chances to refute or confirm empirically a metaphysical thesis such as causal realism.
[36] See Mellor (1974).
[37] It is contrary to Popper (1957), but in line with the theories of Martin (1994), Mellor (1974), Mumford (1998).

independently of the measurements that we decide to make on them. [38] I leave it open, however, whether in addition the combined system of both particles can be ascribed its own propensities – i.e. propensities that neither reduce to, nor supervene on, the properties of the individual particles. (This will turn out to be a useful assumption in some of the common cause models). If it can, then these propensities too will satisfy the condition above and will be possessed even in the absence of any experimental arrangements.

I will be making no further assumptions regarding the nature or correct analysis of propensities. In particular, I will commit myself to neither a realist account in terms of natures or essences, nor an empiricist analysis in terms of conditionals with test-conditions for their manifestations in the antecedents; instead I'll freely help myself to either of these views whenever I find them suitable.[39]

6.1. Propensities for direct-causes

In section 5.1. I defended three alternative lines of defence, or interpretations, of a direct cause model between s_1 and x in EPR. The direct causal influence (i) might not have a material basis; (ii) it might be non-relativistic; or (iii) it might travel backwards in time – in some frame of reference. Each of these interpretations suggests a different understanding of the propensities that may underlie the experiment.

In the first case, for instance, the causal relation can be understood to be a mere conjunction of true counterfactuals *à la* Lewis. For it is then true that: "had s_1 not been the case, it would not have been the case that x", and also that: "had x not been the case, it would not have been the case that s_2". Even if counterfactual dependence is not generally transitive, it is also the case in the context of this direct cause model that: "had it not been the case that s_1, it would not have been the case that s_2". The (partial) cause s_1 might even have its effect x instantaneously and *at a distance* (as in figure 6). On a propensity-based reading of causal influences, we may then say that the outcome-event s_1 endowes – via event x – particle 2 with a certain propensity that displays itself –when appropriately tested by the appropriate device – in a probability distribution over a range of possible (spacelike related to s_1) outcome-events s_2. Notice then that on this reading the causal influence has the effect of changing some hitherto untested propensities of the distant particle; these propensities then reveal themselves, upon the appropriate test, in a probability distribution. Since the causal influence is here analysed counterfactually it might seem that a conditional analysis of propensities would work best. We would then say that event s_1 'endowes' particle 2 with a propensity to s_2 if and only if: *were particle 2 to undergo a subsequent spin measurement by measuring device A_2 it would yield outcome event s_2 with probability $P_y\ (s_2 / s_1)$*. This conditional statement is true within the direct-cause model represented in figure 2, since event x is a necessary partial cause of the display of the statistics predicted by quantum mechanics.

[38] For an interpretation of quantum mechanics along these lines, see Suárez (2004).
[39] Although the conditional analysis of macroscopic dispositional properties suffers from severe difficulties to do with finks and antidotes, it has a better chance in the case of fundamental, and hence irreducible, dispositional properties – see Bird (this volume).

A realist account of propensities is more appropriate for the two other interpretations of a direct cause model. The second interpretation takes it that the causal influences underlying the correlations have a fundamental physical basis, but this fundamental basis ought not be understood in a spacetime context, or at least not a special relativistic one. Special relativity is here taken as a phenomenological theory only, which does not apply to the underlying causal relations. But if it fails to apply to the causal relations then it is hard to see how or why it would apply to the exercise of the propensities, which are just as 'fundamental'. To be more precise: if the causal relation (whereby the measurement outcome-event s_1 on particle 1 effects a change x in the propensities of particle 2) is not subject to the constraints of special relativity, then it is hard to see why the subsequent manifestation of particle 2's propensities (i.e. the displaying of the right probability distribution over outcome-events s_2 under the right kind of test) should be relativistic. This remains so even if x is a physical event that endowes particle 2 with an irreducible property, realistically construed.

Finally, on the third reading of the direct cause model, relativity applies fully and we are required to accept that the causal influence between s_2 and x travels backwards in some frame(s) of reference. So on this reading the manifestation of s_2's subsequent propensity must itself be relativistic, i.e Lorentz invariant. Would this be a problem? I don't think it would, as long as we are referring to the process whereby particle 2's propensities are exercised in s_2 in order to display the correct probability distribution. For the manifestation of the propensity is instantaneous at s_2. It would be a completely different thing if the propensity was, like the causal connection between s_1 and s_2 exercised *at a distance*. But I can see no reason to insist on particle 1 having propensities to yield outcome-events on measurements on particle 2. There is just the need for a causal connection whereby event s_1 can change the propensities of particle 2 at event x, but this is a casual connection between occurrent events, and need not be understood as the exercise of any dispositional property.

6.2. Propensities and common causes

In section 5.2 I presented six different sketches for a common cause model of the EPR correlations. Let me now discuss how best to understand them by means of quantum propensities. One general feature that many of these models share is the following: it will be natural to ascribe the propensities to a larger system that comprises both particles, *in addition to* or *in place of* the propensities of the individual particles. This is most clear in the model developed in section 5.2.3 (figure 8). In this model a hypersurface (or part thereof) that comprises both particles after they have been ejected from the source is the common cause of both outcome-events s_1 and s_2. The propensities responsible for the probability distribution displayed must then be properties possessed by the whole hypersurface (or part thereof). It is in reaction to these propensities of the hypersurface that the right probability distributions over s_1 and s_2 will be displayed.

The model described in section 5.2.1 presents us with a very different account where the common cause is a discrete event (the emission event at the source). Often however this event is represented as the quantum state of the particle pair at the time of their emission, ψ. So if we are going to suppose that the distribution over the outcome events is a display of underlying propensities at the time of emission, we

seem forced to suppose that these propensities are properties of the combined system. But one advantage of the model in 5.2.1. is that, unlike the model considered previously in figure 8, the propensities so ascribed will be possessed by the particle-pair as they are generated at the source, hence they are located at one particular point of spacetime – which is arguably the standard view of propensities and dispositional properties. Moreover, the notion of propensitiy *per se* draws no conceptual distinction between propensities of a common cause that underlie conjunctive forks and those that underlie interactive forks, since the only requirement for their possession – that they both lay out the appropriate probability distributions over their effects – is satisfied in both cases. Thus the propensitiy picture *prima facie* supports critics of the conjunctive fork criterion.

There is, however, a counterintuitive feature of the propensities invoked to explain the model in 5.2.1. These propensities manifest themselves across a spacetime gap, as it were: it is the propensities of the whole combined system *at the time of emission* that manifest themselves later on as the probability distributions over outcome-events *at the time of measurement*. This is a bizarre but apparently coherent kind of what we could call "nonlocal propensity": a dispositional property of a quantum entity placed in some spacetime location that can manifest itself at a different spacetime location.[40] In order to avoid this bizarre feature, one might be tempted to postulate some kind of "*propensity* transmission": each particle transmits along its trajectory the propensities of the particle pair at emission time. Hence the spacetime location of each particle at each instant instantiates the individual particle's properties at that time, but also some of the properties of the combined system at previous times. But this is even more bizarre – if not plainly incoherent: an entity at some precise spacetime location is endowed with propensities belonging to different entities in different spacetime locations. In either case, we seem to be bringing the unavoidable non-locality of quantum mechanics into the description of the operation of quantum propensities. [41]

The same counterintuitive feature of propensities appears again in the model depicted in figure 7 (section 5.2.2) – and this time the resort to the transmission mechanism is not even available. The common cause *d* is the exercise of a propensity that is manifested at *a*, *b*, s_1, and s_2. But there is no material transmission at all between the cause and any of these events. The particle-pair system at the time of emission has certain nonlocal propensities that get actualised at both wings of the experiment, *across a spacetime gap*. This deeply counterintuitive feature would reappear in a propensity based analysis of the model described in section 5.2.4, and might render implausible an analysis of these two models in terms of propensities.

A way to avoid these problems is to allow that in each of the models presented in 2.5.1. and 2.5.2. there is a combination of causal relations between occurrent

[40] For a discussion of the nature of non-locality in quantum mechanics and its relation to causation, see e.g. Berkovitz (2000)

[41] Note, incidentally, that this type of time-lapse between a dispositional property and its manifestation is different from the one discussed by Bird (this volume). According to Bird, at least for macroscopic dispositional properties, the interaction of an experimental device with the propensity in order to yield its manifestation will always take some time – and this gives rise to the possibility of finks and antidotes operating within the interval. But the possibility that I am introducing here is different and unique to quantum mechanics. The interaction of the particles with the experimental device could be instantaneous and take an infinitesimal amount of time, yet there would still be a lapse of (space-)time between the possession of the propensity by the combined system and its manifestation at either wing.

properties and manifestations of underlying propensities – of very much the same kind that I have applied to the direct-cause model. The arrows out of the common cause need not be taken as the manifestation of its propensities. They could instead be interpreted as causal influences whereby some occurrent event (*c* or *d*) causes a change in the propensities of each of the particles.

The backwards causation models of Price and De Beauregard described in section 5.2.5 (depicted in figure 9) could make good use of such combination. In these models there is a backwards causal influence whereby the setting-events in the wings partially determine the properties of the particle pair – including presumably its propensities – at the (earlier) emission event at the source. However, all propensities, unlike some of the causes, are forward looking since the possession of the propensity precedes its manifestation. And moreover they are not "non-local", or "bizarre", in the sense that they manifest themselves instantaneously at their precise location when interacted with by an appropriate measuring device. There are no *spacetime gaps* between possession and manifestation. From the point of view of the nature of the propensities required to make sense of a causal model this might be the most satisfying common-cause model of all!

Finally models that violate hidden locality, such as the model described in section 5.2.6 (depicted in figure 10), appeal to a combination of all the features that I have been discussing here. The causal influence between the common cause *c* and the effects s_1, and s_2 can be understood in terms of propensities of the individual particles carried by the particles themselves – there is no need here to postulate either propensities of the whole composite, nor propensities that can be manifested across spacetime gaps. However, the causal influence of the distant setting events *a* and *b* directly on the effects s_1, and s_2 (if it is to be understood at all as propensities of the measuring devices A and B to generate probability distributions over distant wing outcomes) involve the problematic assumption of manifestations of a propensity across a spacetime gap, characteristic of "non-local" propensities. Since these events *a*, *b* are direct causes of s_1 and s_2, they cannot be understood as causal influences upon the propensities of partricles "1" and "2", but must be understood as exercising their 'powers' directly at a distance.

7. Conclusions

In this paper I have taken on, in a preliminary but I hope suggestive enough form, what I take to be two of the most important unresolved issues in the philosophy of quantum mechanics: What kind of causal models can be given for the EPR correlations? What is the nature of quantum propensities? I have argued that Van Fraassen's 1982 paper constitutes an excellent guide to the first question, and I have sketched some ways in which the answers to the first question might then serve to illuminate the second.

Acknowledgements:

This paper is a improved version of part of chapter 4 of my PhD thesis (Suárez, 1997), and takes further some themes that I also explore in a couple of articles published in

Spanish (Suárez, 1994 and forthcoming). I am grateful for comments on different parts of this paper to audiences in Cambridge, St. Andrews, Oxford, Firenze, London (LSE), Barcelona and Sydney. Research towards the last stages in the writing of this paper has been supported by Spanish Ministry of Science and Technology Project BFF2002-01552: "Causality, Determinism and Probability in Quantum Mechanics and Relativity Theory".

References:

Berkovitz, J. (2000), "The Nature of Causation in Quantum Phenomena", in M. Suárez (ed.) (2000b), *Causality in Physics*, pp. 87-122.

Bird, A. (2004), "Antidotes all the Way Down", this volume.

Bohm, D. (1951), *Quantum Theory*, Prentice Hall.

Bohm, D. (1952), "A Suggested Interpretation of the Quantum Theory in Terms of Hidden Variables, I and II", *Physical Review*, **85**, pp. 166-193. Reprinted in Wheeler and Zurek (1985), pp. 369-396.

Butterfield, J. (1989), "A Space-Time Approach to the Bell Inequality", in Cushing and McMullin, eds., (1989), pp. 114-144.

Cartwright, N. (1988), "How to Tell a Common Cause: Generalizations of the Conjunctive Fork Criterion", in Fetzer, J. ed. (1988), pp. 181-188.

Cartwright, N. (1989), *Nature's Capacities and their Measurement*, Oxford University Press.

Chang, H. and N. Cartwright (1993), "Causality and Realism in the EPR Experiment", *Erkenntnis*, **38**, pp. 169-189.

Costa de Beauregard, O., (1977), "Time Symmetry and the Einstein Paradox", *Il Nuovo Cimento*, **42B**, pp. 41-64.

Cushing, J., (1995), *Quantum Mechanics: Historical Contingency and the Copenhagen Hegemony*, University of Chicago Press.

Cushing and McMullin (eds.), (1989), *Philosophical Consequences of Quantum Theory*, Notre Dame Press.

Dickson, M. and R. Clifton (1998), "Lorentz-Invariance in Modal Interpretations", in Dieks, D. and P. Vermaas (1998).

Dieks, D. and P. Vermaas (1998), *The Modal Interpretation of Quantum Mechanics*, Kluwer Academic Publishers.

Dowe, P. (2000), *Physical Causation*, Cambridge University Press.

Dummett, M. (1954), “Can an Effect Precede Its Cause?”, *Proceedings of the Aristotelian Society, Supp. Volume*, **38**, pp. 27-44.

Einstein, A., B. Podolsky and N. Rosen (1935), “Can Quantum Mechanical Description of the World Be Considered Complete?”, *Physical Review*, **47**, pp. 777-780. Reimpreso en Wheeler and Zurek (1985), pp. 138-141.

Fine, A. (1989), “Do Correlations Need to Be Explained”, in Cushing and McMullin (eds.), pp. 175-194.

Fetzer, J. (ed.), (1988), *Probability and Causality*, Reidel.

Hausman, D. and J. Woodward (1999), “Independence, Invariance and the Causal Markov Condition”, *British Journal for the Philosophy of Science*, **50**, 9, pp. 521-583.

Hofer-Szabo, G, M. Redei and L. Szabo (1999), “On Reichenbach’s Common Cause Principle and Reichenbach’s notion of Common Cause”, *British Journal for the Philosophy of Science*, **50**, 3, pp. 377-399.

Horwich, P. (1987), *Asymmetries in Time*, MIT Press.

Lewis, D. (1986), *On the Plurality of Worlds*, Cambridge University Press.

Martin, C. (1994), “Dispositions and Conditionals”, *Philosophical Quarterly*, **44**, pp. 1-8.

Maudlin, T. (1995), *Quantum Non-Locality and Relativity*, Oxford Blackwells.

Mellor, H. (1974), *The Matter of Chance*, Cambridge University Press.

Mumford, S. (1998), *Dispositions*, Oxford University Press.

Price, H. (1996), *Time’s Arrow and Archimedes’ Point*, Oxford University Press.

Psillos, S. (2002), *Caustion and Explanation*, Acumen.

Reichenbach, H., (1956), *The Direction of Time*, University of California Press.

Salmon, W. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton University Press.

Salmon, W. (1994), “Causality without Counterfactuals”, *Philosophy of Science*, **61**, pp. 297-312.

Schrödinger, E., (1933), “The Present Situation in Quantum Mechanics: A Translation of Schrödinger’s ‘Cat Paradox’ Paper”, Proceedings of the American Philosophical Soceity, **124**, pp. 323-38, 1980. Reprinted in Wheeler and Zurek (1985).

Suárez, M. (1994), “La Cuestión de la Causalidad en el Experimento de Einstein, Podolsky y Rosen: Una Visión Crítica”, Actas del II Congreso de la Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España, UNED.

Suárez, M. (1997), *Models of the World, Data-Models and the Practice of Science: The Semantics of Quantum Theory*, PhD dissertation, London School of Economics.

Suárez, M. (2000a), “The Many Faces of Non-Locality: Dickson on the Quantum Correlations”, *British Journal for the Philosophy of Science*, **51**, 4, pp. 882-892.

Suárez, M. (ed.) (2000b), *Causality in Physics*, special monographic section of the journal *Theoria*, vol. **15**, 37, pp. 5-128.

Suárez(2000c), “Presentation”, in M.Suárez (ed.) (2000b), *Causality in Physics*.

Suárez, M. (2004), “Quantum Selections, Propensities and the Problem of Measurement”, *British Journal for the Philosophy of Science*, **55**, 2, pp. 219-55.

Van Fraassen, B. (1982), “The Charybdis of Realism: Epistemological Implications of Bell’s Inequality”, *Synthese*, 53, pp. 25-38. Reimpreso con correcciones en Cushing and McMullin, eds. (1989).

Wheeler, J. and W. Zurek (1985), *Quantum Theory and Measurement*, Princeton University Press.