Studies in History and Philosophy of Modern Physics 49 (2015) 1–9

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

Retrocausal models for EPR

Richard Corry

University of Tasmania, Australia

ARTICLE INFO

Article history:
Received 23 April 2014
Received in revised form
17 November 2014
Accepted 24 November 2014
Available online 24 December 2014

Keywords:
Retrocausality
EPR
Toy model
Quantum mechanics
Dispositions

ABSTRACT

This paper takes up Huw Price's challenge to develop a retrocausal toy model of the Bell-EPR experiment. I develop three such models which show that a consistent, local, hidden-variables interpretation of the EPR experiment is indeed possible, and which give a feel for the kind of retrocausation involved. The first of the models also makes clear a problematic feature of retrocausation: it seems that we cannot interpret the hidden elements of reality in a retrocausal model as possessing determinate dispositions to affect the outcome of experiments. This is a feature which Price has embraced, but Gordon Belot has argued that this feature renders retrocausal interpretations "unsuitable for formal development", and the lack of such determinate dispositions threatens to undermine the motivation for hidden-variables interpretations in the first place. But Price and Belot are both too quick in their assessment. I show that determinate dispositions are indeed consistent with retrocausation. What is more, I show that the ontological economy allowed by retrocausation holds out the promise of a classical understanding of spin and polarization.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Modern Physics*

1. Introduction

One of the most troubling features of the Copenhagen interpretation of quantum mechanics is its assertion that some measurable properties do not have determinate values before they are measured. This indeterminateness is not only counterintuitive, it is also what makes the measurement problem so difficult for the Copenhagen Interpretation. To account for the fact that measurements always seem to have a determinate outcome, the Copenhagen interpretation posits a special role for measurement in the dynamics of a system—to "collapse" the wavefunction such that the measured property takes on a determinate value. The problem is that there is no clear definition of what counts as a measurement, nor is there any explanation of why measurement should play this role. Hidden-variables interpretations seek to avoid these problems by insisting that the indeterminateness of quantum mechanics is merely epistemic. These interpretations claim that the formalism of quantum mechanics is incomplete in the sense that there are determinate "elements of physical reality" (to use Einstein, Podolsky, & Rosen, 1935 term) that have no counterpart in the formalism. The thought is that many of the puzzling aspects of quantum mechanics might arise from our ignorance of these "hidden" elements of reality. In particular, if the outcomes of all measurements are determined by such elements of reality, then we can interpret the collapse of the wavefunction as an epistemic issue; it represents an updating of our incomplete information about the world rather than a real change in the world from an indeterminate to a determinate state. In this way, hidden-variables interpretations hope to avoid ascribing indeterminateness to the world, and thereby hope to dissolve the measurement problem.

However, hidden-variables interpretations face a major problem: there are a number of No Hidden Variables theorems which seem to show that the assumption of hidden variables is incompatible with quantum mechanics. Of particular interest is Bell (1964) variation of the "EPR" thought experiment first presented in 1935 by Albert Einstein, Boris Podolsky, and Nathan Rosen (ironically, Einstein, Podolsky, and Rosen presented the original version as an argument that quantum mechanics must be incomplete). Bell showed that in this thought-experiment, hidden-variables interpretations make predictions that are at odds with accepted quantum mechanics. Later experimental results seem to uphold the predictions of quantum mechanics rather than hidden-variables (Aspect, Dalibard, & Roger, 1982).

Physicists were quick to note that hidden-variables interpretations could be saved if they allow non-local interactions; in particular, if the interpretations include instantaneous action at a distance then they would not be inconsistent with quantum

E-mail address: Richard.Corry@utas.edu.au

http://dx.doi.org/10.1016/j.shpsb.2014.11.001
1355-2198/

mechanics. Thus the conclusion was drawn that hidden-variables theories must be non-local. Physicists have long been suspicious of action-at-a-distance (despite such action playing a central role in Newton's theory of gravitation), and *instantaneous* action-at-a-distance is particularly problematic since it assumes an objective notion of simultaneity, in conflict with Einstein's theory of relativity. For this reason the EPR thought experiment is regarded by many as an important nail in the coffin of hidden-variables interpretations of quantum mechanics. This conclusion is a little unfair, however, since the EPR experiment also shows that non-hidden-variables interpretations like the Copenhagen interpretation must likewise involve a kind of instantaneous action-at-a-distance (this was essentially the point of Einstein, Podolsky, and Rosen's paper) and are thus no better off than hidden-variables theories in this respect.

There is, however, a way to make a local hidden-variables interpretation compatible with quantum mechanics. Bell's argument assumes that the values of the hidden variables are independent of the future settings of the experimental apparatus. Indeed, this same assumption is made in all No Hidden Variables theorems. Thus, a number of writers have suggested that we can resolve many of the puzzles of quantum mechanics if we allow the possibility of *retrocausation*, whereby the properties of a system can be influenced by future events (see, e.g. Costa de Beauregard, 1976; Cramer, 1980; Dowe, 1997; Miller, 1996; Price, 1997; Sutherland, 1983; Wharton, 2007).

Of course, retrocausation is itself rather counterintuitive and is often dismissed as involving paradox or problems for free will (see, for example, Bell's comments in Davies & Brown, 1986, pp. 49–50). In response to such attitudes, Price introduced the strategy of investigating retrocausation by constructing "toy models" that can be used to explore and elucidate the possibilities of retrocausation. The first of these toy models—the *Helsinki model*—is designed to represent some very general features of retrocausation, and he expresses his hope that further models will be developed which capture more specifically quantum phenomena. In particular, he comments that a model that includes Bell-like correlations is the "retrocausal toy modeller's Holy Grail" Price (2008, p. 761).

This paper takes up Price's challenge and develops a retrocausal toy model of the Bell-EPR experiment. The model shows that a consistent, local, hidden-variables interpretation of the EPR experiment is indeed possible, and gives a feel for the kind of retrocausation involved. However, the model also makes clear a problematic feature of retrocausation: it seems that we cannot interpret the hidden elements of reality in a retrocausal model as possessing determinate dispositions to affect the outcome of experiments. This is a feature which Price (1997, p. 250) has embraced, however Gordon Belot has argued that this feature renders retrocausal interpretations "unsuitable for formal development" (Belot, 1998, p. 479), and the lack of such determinate dispositions threatens to undermine the motivation for hidden-variables interpretations in the first place. But Price and Belot are both too quick in their assessment. I will show that the retrocausal model is consistent with determinate dispositions so long as one accepts a particular view of the metaphysics of dispositions. I will also consider two variations of the original retrocausal model which allow for determinate dispositions even without this metaphysical assumption.

2. Modeling EPR

In the original EPR thought experiment, two particles are created in an entangled state, the two particles are then separated, and a measurement is performed on each of the separated particles. In what follows I will focus on Bohm's (1951) variation of the thought experiment. In this variation, we can set each of our measuring devices to make one of three different measurements, and each measurement gives one of two possible results (for example we might set the devices to measure the spin along three different axes). Call these three settings *A*, *B*, and *C*. The original EPR thought experiment is recovered if we allow only two settings and consider only situations in which the same setting is chosen for the measurement of each particle. For convenience I will refer to Bohm-type EPR experiments simply as EPR experiments from now on.

The interesting features of the thought experiment derive from the facts that (i) the two particles are in an entangled state, meaning that they are not independent in some sense (to be discussed below); and (ii) measurements *A*, *B*, and *C* measure properties that are not simultaneously given determinate values in any quantum state description.

Following Price (2008), the models presented below focus on the causal structure of the thought experiment, and the relevant facts about entanglement and measurement are represented by placing constraints on the possible interactions. We begin, then, by noting that the EPR experiment involves three interactions: an interaction that produces a pair of entangled systems, and two measurement interactions. This structure is depicted in Fig. 1.

Here X is the device that produces the entangled state, S_1 and S_2 are the settings of the two measuring devices while O_1 and O_2 are the observed outcomes of the two experiments. We will let S_1 and S_2 each take values from the set $\{A, B, C\}$, representing the three measurement settings, while O_1 and O_2 each take values from the set $\{+, -\}$, to represent the two possible measurement outcomes. We will leave the possible values of X unspecified.

Vertical separation between nodes represents temporal separation, and we will stipulate that the future is towards the top of the page. Horizontal separation between nodes represents spatial separation. The unlabeled internal paths represent two physical systems that interact at one point in time, then move away from each other. Each of these systems is then involved in a measurement interaction at some later point in time. For convenience sake, I will refer to the leftmost of these systems as "particle 1" and the rightmost as "particle 2". In general, however, these systems need not be thought of as particles. We will ensure that our models are consistent with special relativity by insisting that all paths be null or timelike (and hence cannot represent systems traveling faster than the speed of light). Finally, let us stipulate that the two measurement interactions are simultaneous in the laboratory rest frame. Note that this last stipulation together with the restriction that paths be null or timelike imply that there can be no path directly connecting the two measurement interactions.

If we assume that at X we have a device for producing two particles in an appropriately entangled state, then quantum mechanics predicts—and experiment seems to confirm (Aspect et al., 1982)—the following two facts:

Fact 1: Whenever S_1 and S_2 have the same setting (regardless of what it is), O_1 and O_2 have opposite outcomes. So graphs

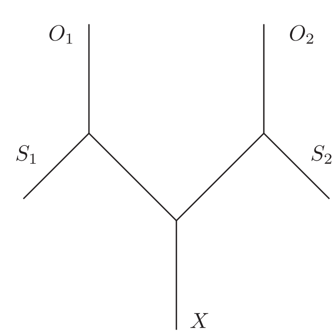

Fig. 1. Causal model of the EPR experiment.