

Lessons of Bell's Theorem:  
Nonlocality, yes; Action at a distance, not  
necessarily.

Wayne C. Myrvold  
Department of Philosophy  
The University of Western Ontario

Forthcoming in Shan Gao and Mary Bell, eds., *Quantum Nonlocality and  
Reality – 50 Years of Bell's Theorem* (Cambridge University Press)



# Contents

1	Introduction.	<i>page</i> 1
2	Does relativity preclude action at a distance?	2
3	Locally explicable correlations.	5
4	Correlations that are not locally explicable	9
5	Bell and Local Causality	11
6	Quantum state evolution	13
7	Local beables for relativistic collapse theories	17
8	A comment on Everettian theories	19
9	Conclusion	20
10	Acknowledgments	20
11	Appendix	20
	<i>References</i>	25



## 1 Introduction.

Fifty years after the publication of Bell's theorem, there remains some controversy regarding what the theorem is telling us about quantum mechanics, and what the experimental violations of Bell inequalities are telling us about the world. This chapter represents my best attempt to be clear about what I think the lessons are. In brief: there is some sort of nonlocality inherent in any quantum theory, and, moreover, in any theory that reproduces, even approximately, the quantum probabilities for the outcomes of experiments. But not all forms of nonlocality are the same; there is a distinction to be made between action at a distance and other forms of nonlocality, and I will argue that the nonlocality required to violate the Bell inequalities need not involve action at a distance. Furthermore, the distinction between forms of nonlocality makes a difference when it comes to compatibility with relativistic causal structure.

The Bell locality condition is a condition which, if satisfied, renders possible a completely local account of the correlations between outcomes of spatially separated experiments. Bell's theorem tells us that the probabilities that we compute from quantum mechanics do not admit of such an account. Modulo auxiliary assumptions that, in my opinion, should be noncontroversial, this yields the conclusion that there must be something nonlocal about any theory that yields the quantum mechanical probabilities or anything close to them.

One way to violate the Bell locality condition—and this is the way that any deterministic theory must do it—is via straightforward action at a distance. However, we should also consider the possibility of a stochastic theory with probabilistic laws that involve irreducible correlations. The sort of nonlocality inherent in such a theory involves relations between events that are significantly different from causal relations as usually conceived. Perhaps surprisingly, such a theory, unlike its deterministic cousins, need not invoke a distinguished relation of distant simultaneity.

This makes a difference for the situation of the two sorts of theories *vis-a-vis* relativity. Though a deterministic theory can achieve an *appearance* of being relativistic, at the level of observable phenomena, it can do so only by introducing a preferred foliation and with it a causal structure at odds with relativistic causal structures. Things are otherwise with dynamical collapse theories.

Conclusions along these lines have, of course, been drawn by many others before, though not everyone is convinced. Travis Norsen (2009), in particular, has expressed vigorous skepticism. I believe, with Ghirardi (2012), that,

though Norsen’s criticisms of some earlier arguments are well taken, there is still a case to be made. My own understanding has been strongly influenced by the work of my teacher, Abner Shimony (1978, 1984, 1986), but mention should also be made of Jarrett (1984), Skyrms (1984), Redhead (1987), Ghirardi and Grassi (1996), and Ghirardi (2012)—among others! However, the arguments presented here, though inspired by and indebted to all of the above-mentioned authors, differs in some details from each of them.

## 2 Does relativity preclude action at a distance?

One reason for making a distinction between action at a distance and other forms of nonlocality has to do with compatibility with relativistic spacetime structure. It is often said that relativity precludes action at a distance. This is right, I think. But, in order to get clear about what is precluded by what, a few words are in order about the notion of causation, and about what I mean by a relativistic spacetime.

### 2.1 Causation

A paradigm case of causation is an intervention and its causal consequences. In the most straightforward cases, there is some variable  $X$ , whose values are ones that some agent could, at least in principle, choose between, holding fixed the background state of other relevant variables, such that the probability distribution for some other variable  $Y$  depends on the value of  $X$ . It is typically assumed that the realization of  $Y$  occurs at a later time than the setting of  $X$ . If we let  $\Lambda$  be the background variables (besides  $X$ ) relevant to  $Y$ , then, relative to a particular specification  $\Lambda = \lambda$ ,  $X$  is causally relevant to  $Y$  if and only if, for some  $y$ , and some possible values  $x, x'$  of  $X$ ,

$$Pr(Y = y|X = x, \Lambda = \lambda) \neq Pr(Y = y|X = x', \Lambda = \lambda).$$

We will say that the choice of  $X$  has a causal bearing on  $Y$ , relative to background conditions  $\lambda$ .

In the most straightforward cases, the variable that is the target of intervention is something that some agent, human or otherwise, could in principle set. But the notion of agency is not essential. When talking about causation, I will, however, assume that it makes sense to consider some process that fixes the value of  $X$  without changing the other parameters relevant to  $Y$ , and to consider the result of such an intervention on the probability of  $Y$ . This is what distinguishes a causal relation between  $X$  and  $Y$  and situations in which  $X$  is merely informative about  $Y$ , without being causally

relevant to it. To invoke the standard example, the reading of a barometer gauge may be informative about a coming storm (because it is informative about a drop on pressure which *is* causally relevant to the storm), but it is not causally relevant, because an intervention that changes the barometer reading (perhaps via a localized change in atmospheric pressure) without changing such things as large-scale atmosphere pressure doesn't affect the probability of  $Y$  but only disrupts the informational link between  $X$  and  $Y$ .<sup>1</sup>

There are complications, familiar from the literature on causation, having to do with matters of overdetermination, pre-emption, and the like. But these are not, as far as I can see, important for the issue at hand, and maybe safely disregarded for present purposes.

Signals are special cases of cause-effect relations; the sender chooses between various setting of the signalling device, and the signal obtained by the receiver is informative about the sender's choice. But there may be causal relations that don't permit signalling. It might be the case that there are other factors,  $Z$ , imperfectly known to sender and receiver, such that, for any value of  $Z$ , the probability of  $Y$  depends on  $X$ , but it does so in different ways for different values of  $Z$ . Suppose, for example, that Alice can choose between pushing two buttons, labelled 0 and 1, and, subsequently, Bob will see a light bulb light up either red or green. But the relation between the button Alice pushes and the colour of the light Bob sees depends on the setting of a hidden switch; when the switch is set one way, button 0 yields a red light and button 1 a green, but when it is set another way, the relation between the button pushed and the colour of the light is reversed. The setting of the switch is hidden from Alice and Bob and fluctuates unpredictably. In such a case, Alice's choice of which button to push has a causal effect on the colour of the light that Bob sees, but *which* causal effect it has is unknown to Alice and Bob (This is analogous to the cause-effect relations between the two sides of an EPR-Bohm experiment found in the de Broglie-Bohm theory; for any specification of initial positions of the two particles, the experimental setting of the earlier experiment affects the result of the other experiment, but this relation cannot be used for signalling because neither Alice nor Bob has epistemic access to these initial positions).

Now, if, at the time of the setting of  $X$ , the realization of  $Y$  has already occurred—that is,  $Y$  has already taken on some definite value—it is hard to see how to make sense of probabilities for  $Y$  taking on this or that value, and hence, difficult to make sense of the choice of  $X$  having an effect on

<sup>1</sup> See Woodward (2013) and references therein.

the probability of  $Y$  taking on this or that value. In my view (admittedly not shared by all), these difficulties are insurmountable, and I adopt the usual conception of causation on which it involves an inherently temporally asymmetric relation. In some circles, retrocausality, or backwards-in-time causation, is invoked as a serious possibility, and one that, moreover, is meant to be relevant to quantum nonlocality.<sup>2</sup> Assessment of such proposals is beyond the scope of this paper.

The notion of cause that will be operative in this chapter is temporally asymmetric; the cause must precede the effect. On such a notion, the relation of temporal precedence constrains the relation of potential causal influence; an event  $x$  is a potential causal influence on  $y$  only if it is in the past of  $y$ . It is this feature of causation that precludes causal influences between spacelike separated events in a relativistic spacetime.

#### *Relativistic spacetimes*

From the foregoing discussion, we take it that an event  $x$  is a potential cause of another event  $y$  only if  $x$  is in the temporal past of  $y$ . There seems to be no harm in widening this to “if and only if,” in which case the relations of temporal precedence and potential causal influence coincide. In Galilean spacetime, the spacetime manifold is partitioned into equivalence classes of simultaneity, and these equivalence classes are totally ordered by the relation of temporal precedence. Any event is either in the past of, in the future of, or simultaneous with any other. This is a marked—and to my mind, the most interesting—difference between Galilean spacetime and Minkowski spacetime. In Minkowski spacetime, the relation of temporal precedence is given by the light cone structure. Though the  $t$ -coordinate of any Lorentz frame partitions Minkowski spacetime into equal- $t$  equivalence classes, none of these captures the relation of temporal precedence, as there are always events that are ordered by a given reference frames’s  $t$ -coordinates but are nonetheless spacelike separated, so that neither temporally precedes the other.

To speak more generally: suppose we have a spacetime with a relation  $\ll$  of temporal precedence. Take this to be transitive and antisymmetric (no temporal loops).

$$\begin{aligned} (x \ll y \ \& \ y \ll z) &\Rightarrow x \ll z. \\ x \ll y &\Rightarrow \neg(y \ll x). \end{aligned} \tag{1}$$

Define the relation  $\sim$  as the relation that holds between two events when

<sup>2</sup> See Faye (2010) for an overview; for retrocausality as a source of quantum nonlocality, see Evans et al. (2013) and references therein.



neither temporally precedes the other.

$$x \sim y \equiv_{df.} \neg(x \ll y \vee y \ll x). \quad (2)$$

This relation is symmetric by construction and, by the antisymmetry of  $\ll$ , reflexive. It is, therefore, an equivalence relation if and only if it is transitive. On Galilean spacetime, the relation  $\sim$  defined by the relation (2) is a transitive relation and therefore partitions the space time into equivalence classes of simultaneous events. In Minkowski spacetime however, the relation of temporal inconnectability is not transitive.

By a relativistic spacetime I will mean a spacetime endowed with a causal structure such that for any  $x, y$  such that  $x \sim y$ , there exists  $z$  such that  $x \sim z$  and  $z \ll y$ .

Now, we could have a theory that is formulated against a background of Minkowski spacetime but is such that the causal structure depends on the distribution of matter or some other fields. One might even have it that there is a field that, for appropriate configurations of the field, yields a partition of the spacetime into hypersurfaces of simultaneity that are made use of by the dynamics of the matter distribution, in such a way that these hypersurfaces yield the relation of causal connectibility. In such a case, I will say that the physical relation of causal connectibility, the one that counts, is the introduced relation, not the one given by the background spacetime.<sup>3</sup> In such a situation, the causal structure is not relativistic.

### 3 Locally explicable correlations.

In a classical setting, it is easy to produce set-ups that result in correlations (or anticorrelations) between distant events. Flip a fair coin, and, depending on the outcome, put a ball in one of two boxes,  $A$  and  $B$ . Send one of these boxes to Paris, the other to Tokyo. The set up yields probability of  $1/2$  that there is a ball in box  $A$ , probability  $1/2$  that there is a ball in box  $B$ , but probability  $0$  that there is a ball in both boxes.

In this set-up, the outcomes of the experiments in Paris and Tokyo, which consist of opening the boxes and looking at them, are determined by initial conditions at the source. But this isn't a necessary condition, in order for there to be distant correlations. Suppose we have two coins, one with a bias

<sup>3</sup> This is what is done, for instance, by Dürr et al. (1999), who formulate a Bohmian theory against a background of Minkowski spacetime by introducing an auxiliary field that picks out a distinguished foliation that is then used to formulate the dynamics of the theory. In this theory, there are causal relations between events that cannot be connected by a light signal; what counts, for the dynamics, is temporal precedence according to the introduced foliation. Though formulated in terms of Lorentz-covariant equations, this is not a theory whose causal relations are relativistic in the sense used here. See Dürr et al. (2014) for discussion.

of 2 to 1 in favour of Heads, the other with the opposite bias. We put one coin in each box, in such a way that each box has an equal chance of getting the Heads-biased coin. The boxes are sent to Paris and Tokyo, where they are opened and the coins are flipped.

This scenario yields a probability of  $1/2$  that the coin in Tokyo will land Heads, since the coin is biased either 2 to 1 towards Heads or 2 to 1 against, with equal chance of each. Similarly, the probability is  $1/2$  that the coin in Paris will land Heads. But these two probabilities are not independent. Since the two tosses are independent tosses of two coins, one with probability  $2/3$  of Heads, the other with probability  $1/3$ , the probability that they will both land Heads is  $2/9$ , which is less than  $1/4$ .

The correlation, or rather, anticorrelation, between the results in Tokyo and Paris is completely explicable in local terms. The probabilistic dependence between the distant outcomes isn't due either to action at a distance or to any irreducible probabilistic dependence between distant events.

What the two scenarios have in common is that a choice is made at the source between probability distributions, each of which is one on which the outcomes of distant experiments are probabilistically independent. The resulting probability distribution for the outcomes of the distant experiments is a mixture of distributions on which the outcomes are independent. Call such correlations *locally explicable*.

Now, consider a set-up such as the one envisaged by Bell. Two systems are prepared at some source, and sent to distant locations  $A$ ,  $B$ , where there is a choice of experiments to be formed. Let  $\lambda$  be a specification of local initial conditions at the source relevant to the outcomes. We assume that, for any choice  $a, b$ , of experiments at  $A$  and  $B$ , respectively, and any specification of relevant initial conditions  $\lambda$ , that there is a probability distribution  $P_{a,b}(x, y|\lambda)$  over outcomes of the experiments. We also assume that there is a probability distribution over the initial conditions  $\lambda$  given by  $\rho_{a,b}(\lambda)$ , such that

$$P_{a,b}(x, y) = \int d\lambda \rho_{a,b}(\lambda) P_{a,b}(x, y|\lambda). \quad (3)$$

Given  $P_{a,b}(x, y|\lambda)$ , we define marginals

$$\begin{aligned} P_{a,b}^A(x|\lambda) &= \sum_y P_{a,b}(x, y|\lambda) \\ P_{a,b}^B(y|\lambda) &= \sum_x P_{a,b}(x, y|\lambda) \end{aligned} \quad (4)$$

We suppose that it is possible to arrange things so that whatever device it is that switches between alternative experiments can be rendered effectively

independent of the distribution  $\rho_{a,b}(\lambda)$  of relevant initial conditions—an assumption implicit in Bell’s original exposition, and made explicit following Bell’s exchange with Shimony, Horne, and Clauser (Shimony et al., 1976; Bell, 1977). The preparation of the systems and the switching events will, of course, have events in their common past, but we assume that these can be effectively screened off. This assumption may be called the “Free Will” assumption, as long as one remembers that it is so-called with tongue in cheek; metaphysical issues concerning the free will of the experimenters are not at stake, but only the more prosaic assumption that it is possible to set up things so that there is effective independence of state preparation and experiment subsequently performed, an assumption so pervasive that it is difficult to see how we could engage in experimental science without it.<sup>4</sup>

In accordance with this, we assume that the switching is done in a way that the distribution of relevant initial conditions  $\lambda$  is effectively independent of the settings of the apparatus, and write

$$\rho_{a,b}(\lambda) = \rho(\lambda). \quad (5)$$

Now, because the experimental settings are the sorts of things that could, in principle, be manipulated in a systematic way, then, if the marginals at  $B$  depend on the setting at  $A$ , this is clearly a case of causal influence. The condition that the setting at  $A$  have no causal influence on the result at  $B$ , and *vice versa*, is the condition that Jarrett called *Locality* and Shimony called *Parameter Independence*. This is the condition that, for all possible outcomes  $(x, y)$  of the two experiments, and for all settings  $a, a'$  and  $b, b'$ , and all  $\lambda$ ,

$$P_{a,b}^A(x|\lambda) = P_{a,b'}^A(x|\lambda) \quad (6)$$

$$P_{a,b}^B(y|\lambda) = P_{a',b}^B(y|\lambda)$$

When this holds, we will drop the redundant parameters and write

$$P_{a,b}^A(x|\lambda) = P_a^A(x|\lambda) \quad (7)$$

$$P_{a,b}^B(y|\lambda) = P_b^B(y|\lambda)$$

Suppose that we are in a relativistic spacetime, and suppose that the two experiments are performed in spacetime regions  $A, B$  such that  $A \sim B$ . Then, because the spacetime is relativistic, there are regions  $X, Y$  such that  $X \sim B, Y \sim A$ , but  $X \ll A$  and  $Y \ll B$ . Suppose that it is possible to arrange things so that the setting of the  $A$ -apparatus occurs in spacetime

<sup>4</sup> See Bell (2004, pp. 102–103) for a lucid discussion.

region  $X$ , and the setting of the  $B$ -apparatus occurs in  $Y$ . In such a case the causal structure of our spacetime *requires* that the conditions (6) be satisfied.

The condition that, for any given settings  $(a, b)$ , any correlations between outcomes be locally explicable, is the condition that the probability distributions be mixtures of distributions on which the outcomes are independent. This is the condition that Jarrett (1984) called *Completeness* and Shimony (1986), *Outcome Independence*.

$$P_{a,b}(x, y|\lambda) = P_{a,b}^A(x|\lambda) P_{a,b}^B(y|\lambda) \quad (8)$$

In a relativistic spacetime, for appropriate experimental arrangements, the condition of Parameter Independence (PI) is the condition that causal relations respect the temporal structure of the spacetime. The condition of Outcome Independence (OI) is the condition any correlations between outcomes be locally explicable. The conjunction of the two is the condition often called the *Bell Locality* condition.

$$P_{a,b}(x, y|\lambda) = P_a^A(x|\lambda) P_b^B(y|\lambda) \quad (9)$$

Now, as is well-known, the Bell Locality condition (9), together with the Free Will assumption, entail the satisfaction of the CHSH inequality, which is violated by quantum-mechanical probabilities. Furthermore, the experimental evidence vindicates the quantum-mechanical probabilities.

A deterministic theory, one on which all probabilities are either zero or one, has to satisfy OI. Thus, if the theory is to produce Bell-Inequality violating statistics, it must violate the condition PI, which, we have argued, is for appropriate set-ups, required by relativistic spacetime structures. Deterministic theories that yield quantum predictions cannot respect relativistic causal structure.

It doesn't follow that quantum mechanical predictions and relativity are incompatible *tout court*. I have argued that PI is required by relativistic causal structure, but not that OI is. In the next section I will argue that, if we are willing to accept a stochastic theory with correlations that are not reducible to or explicable in terms of local factors, we can retain relativistic causality.

## 4 Correlations that are not locally explicable

### 4.1 A toy example

Let us consider another toy example. We have two boxes, *A* and *B*. On each box is a button, a light, and a switch, with two positions labelled 0 and 1. When the button on a box is pushed, its light glows either red or green.

Suppose that the physical laws governing the boxes are chancy; the probabilities do not supervene on any deeper underlying theory. And suppose that, when the button on either box is pushed, there are equal probabilities for the box's light to glow red or green, for any switch settings. Suppose also that, when the switch settings on the boxes agree, the lights on the two boxes always light up the same colour, and that, when the two switch settings disagree, there is a nonzero probability of getting a red light on one box and a green light on the other. By stipulation, these probabilities do not supervene on deeper facts about the boxes; we have given a complete description. And these probabilities, since they are part of chancy physical laws, are to be thought of as objective probabilities.

In this example, there is no causal influence of the switch setting on Box *A* on the outcome of pushing the button on either Box *A* or *B*; the probabilities are  $1/2$ , independent of the switch setting; we have no Parameter Dependence. When the two switch settings agree, the outcomes of pushing the buttons on the two boxes are correlated; there is Outcome Dependence. Another way of saying this is that the conditional probability of one outcome, conditional on the other, is not the same as its unconditional probability. But do we have, in this example, a *causal* influence of one on the other?

This example does not fall within the paradigm case of a possible intervention changing the probability of another event. By hypothesis, the outcomes are chancy, and there is no way to intervene on the system to *set* the outcome to either red or green—that is, no process that reliably produces a red or green result while holding other relevant parameters fixed. But, you may say, perhaps there are cases of causal relations that don't fall within the scope of paradigm cases like this. Fair enough! But we should bear in mind that the example as described is symmetric under interchange of the two boxes, and so there is nothing in the setup to distinguish between cause and effect. One could, of course, expand the use of talk of causal relations to include symmetric relations that don't distinguish between cause and effect. In my opinion this would be an unfortunate terminological choice, but it is nothing more than that. What is important is that, if such a choice were made, we would lose any grounds for thinking that such a relation is prohibited by the temporal structure of a relativistic spacetime. It is only causal

asymmetry, that is, the idea that causation requires temporal precedence, that leads to the conclusion that there can be no causal relations between spacelike separated events in a relativistic spacetime.

If we want to make the example more analogous to quantum mechanics, we can add in an analogue of local state preparation. Suppose that there is another knob on the box that sets the outcome to red or green, but also that (in keeping with the analogy), if we use the knob on Box  $A$  to set its outcome variable to red (or green), there is an equal chance of either outcome for Box  $B$  (and *vice versa*, with the roles of the two boxes switched). In such a case, there would be a possible intervention to set the outcome variable of either box, but such an intervention would have no causal effect on the outcome variable of the other box. A state preparation on one box is like a local intervention on a barometer; it sets the value of the barometer reading at the cost of destroying the informational link between the barometer reading and the distant storm.

Though there is nothing specifically quantum about this example, we do have an interesting form of nonseparability. Let  $A, B$  be variables that specify the switch setting of the two boxes, and let  $X, Y$  be the variables that specify the two outcomes, that is, whether the lights are red or green. Let  $C = \langle A, B \rangle$  be a variable specifying the pair of switch settings, and let  $Z = \langle X, Y \rangle$  specify the pair of outcomes. Neither  $A$  nor  $B$  is causally relevant to  $X$  or  $Y$ . Nor is  $C$ , since the probability of red and green at each end of the experiment is  $1/2$ , regardless of the switch settings. Neither  $A$  nor  $B$ , by itself, is causally relevant to  $Z$ . But  $C$  *is* causally relevant to  $Z$ .

There is in this example no nonlocal causation, because there are no causes with distant effects. But we do have a radical nonseparability of causes;  $C$  is causally relevant to  $Z$ , though its component parts,  $A$  and  $B$ , are not.

#### 4.2 Probability and becoming

Suppose that our toy example lives in a relativistic spacetime, and that the two button-pushings take place at spacelike separation. It might be convenient, when giving an account of events, to pick a foliation and use it to assign global times. The natural thing to do is to assign probabilities to events on a given hypersurface conditional on events to the past of that hypersurface. So, on such an account, prior to  $B$ 's button being pushed, we assign a probability  $1/2$  to red and green for  $B$ -outcomes on hypersurfaces to the past of the  $A$ -experiment, and, for hypersurfaces to the future of the  $A$ -experiment, probabilities conditional on the result of that experiment. So,

for example, when the switch settings agree, the probability for a red light at  $B$ , conditional on the result of the  $A$ -experiment, will be either 1 or 0.

It should be clear that transition from the probability to the conditional probability, different from  $1/2$ , that occurs when we pass from a hypersurface that has the  $A$ -experiment to its future to one with the  $A$ -experiment in its past, is not a change in any intrinsic property of  $B$ . Any temptation to think of it as such results from a lingering assumption of separability. This is worth saying because there is a temptation to say that, when we do the  $A$ -experiment, it instantaneously changes some fact about  $B$ , namely, the probability of glowing red when its button is pushed, and we may wonder when this fact about  $B$  changes, and be tempted to conclude that our toy example can't live comfortably in a relativistic spacetime after all.

Suppose, now, that the  $A$  experiment occurs in the past of the  $B$  experiment. Then, once the outcome of the  $A$ -experiment is in  $B$ 's past light cone, it is to the past of every spacelike hypersurface intersecting  $B$ 's world line, and, by the rule that stipulates that we associate, with any spacelike hypersurface, probabilities conditional on events to the past of that hypersurface, then all hypersurfaces will agree on ascribing the conditional probabilities to  $B$ . Then, and only then, can we regard this conditional probability as a local beable.

## 5 Bell and Local Causality

Two lectures given in 1989, the last full year of his life, shed light on Bell's views on the locality conditions. One, "La Nouvelle Cuisine," (Bell, 1990) was presented in June, in Eindhoven, the other (Bell, 2007) in November, in Trieste.

In the June lecture, Bell formulated a principle that he called the *Principle of Local Causality*. We are first given an informal gloss,

The direct causes (and effects) of events are near by, and even the indirect causes (and effects) are no further away than permitted by the velocity of light (Bell, 2004, p. 239).

Bell followed this by what was intended to be a sharper and cleaner version of this intuitive idea.

A theory is said to be locally causal if the probabilities attached to values of local beables in a spacetime region 1 are unaltered by specification of values of local beables in a space-like separated region 2, when what happens in the backward light cone of 1 is already sufficiently specified, for example by a full specification of local beables in a space-time region 3 [which, in the diagram supplied by Bell, is a cross-section of the backward light cone of region 1] (pp. 239–240).

It seems to me that this condition is strictly stronger than the intuitive condition that motivates it. If Bell's second principle is satisfied, then (as Bell proceeds to argue), all correlations are locally explicable and there is no action at a distance. But the entailment doesn't go the other way, as is illustrated by the toy example of the previous section. That example is compatible with all *causal* relations being local ones; it's just that there are irreducible correlations, not explicable in terms of causal relations, local or nonlocal. In the toy example, the relation between outcomes of the two experiments are nonlocal relations, but they are not nonlocal causal relations, because they are not causal.

Again, it makes a difference because there are two ways to violate the Bell Inequality condition. Theories that violate Parameter Dependence exhibit action at a distance and do not sit well with relativistic causal structures. But stochastic theories with irreducible correlations and no Parameter Dependence need not be at odds with relativistic causal structure.

This is something that, as pointed out by Ghirardi (2012), Bell himself appreciated. In his paper "Are There Quantum Jumps?" (Bell, 1987), Bell showed that the GRW theory exhibits relative time-translation invariance—that is, for entangled pairs of particles that don't interact with each other, the theory does not require a distinguished notion of distant simultaneity, a property that he called "a residue, or at least an analogue, of Lorentz invariance" in the nonrelativistic theory. He ended with the remark,

I am particularly struck by the fact that the model is as Lorentz invariant as it could be in the nonrelativistic version. It takes away the ground of my fear that any exact formulation of quantum mechanics must conflict with fundamental Lorentz invariance (Bell, 2004, p. 209)

This is in contrast to the Bohm theory, which requires a distinguished relation of distant simultaneity for its formulation, and hence cannot be made Lorentz invariant at the fundamental level, though there is hope for a Bohm-like theory whose observable consequences *are* Lorentz invariant, because inaccessibility of the exact particle trajectories leaves the preferred foliation empirically undetectable.

In the Trieste lecture, Bell discussed the prospects for a genuinely relativistic version of a dynamical collapse theory, and concluded that the difficulties encountered by Ghirardi, Grassi, and Pearle in producing a genuinely relativistic version of the Continuous Spontaneous Localization theory (CSL), a theory that would be "Lorentz invariant, not just for all practical purposes but deeply, in the sense of Einstein, eliminating entirely any privileged reference system from the theory" (p. 2931), were "Second-Class Difficulties,"



technical difficulties, and not deep conceptual ones. This has been borne out by the work of Bedingham (2011), who has constructed a relativistic version of the CSL theory; see also Tumulka (2006) for a relativistic version of the GRW theory.

Bell's hope that there could be a deeply Lorentz invariant collapse theory that would not already be refuted by existing experiments indicates that he did not think that the requirement he called *Local Causality* was something required by relativity, since any theory that is locally causal in Bell's sense would have to satisfy the Bell Inequalities.

## 6 Quantum state evolution

### 6.1 Local and nonlocal operations: the nonrelativistic case

We have superluminal causation if an intervention on a variable that is a local beable pertaining to a region  $A$  affects the probability of an event that takes place at spacelike distance, that is, the realization of a variable that is local beable pertaining to a region  $B$ , at spacelike separation from  $A$ . For this to even make sense, there must be some local beables.

When separability obtains—that is, if all beables supervene on local beables—then it is easy to distinguish between local actions and nonlocal actions. Given a composite system  $AB$ , with spatially separated parts  $A$  and  $B$ , if a complete specification of the state of the composite system is determined by local beables pertaining to  $A$  and  $B$ , then we can partition changes to the system into changes pertaining to  $A$  and changes pertaining to  $B$ .

When separability does not obtain, as is the case for quantum systems, things are a bit trickier, as the state of the system is not divisible into a part pertaining to  $A$  and a part pertaining to  $B$ . But this does not mean that we cannot distinguish between local and nonlocal action. Consider a pair of systems,  $A$  and  $B$ , with associated Hilbert spaces  $\mathcal{H}_A$ ,  $\mathcal{H}_B$ . Let  $\mathcal{A}$  and  $\mathcal{B}$  be the algebras of operators on  $\mathcal{H}_A$  and  $\mathcal{H}_B$ . When the evolution is unitary, it is easy to distinguish between evolution that is local, and evolution that involves interaction between two separated systems  $A$  and  $B$ . If the total Hamiltonian is a sum of Hamiltonians  $H_A \in \mathcal{A}$  and  $H_B \in \mathcal{B}$ , then the unitary operator  $U$  that implements the evolution will factor,

$$U = U_A U_B, \tag{10}$$

with  $U_A \in \mathcal{A}$  and  $U_B \in \mathcal{B}$ .

When dynamical collapse is involved, we can say something similar. It is generally taken to be the case that physically realizable state changes to

the system  $A$  must be completely positive mappings of the state space into itself. Any such mapping can be represented by a set  $\{K_i\}$  of operators, such that the density operator  $\rho$  representing the state undergoes the change

$$\rho \rightarrow \sum_i K_i \rho K_i^\dagger, \quad (11)$$

where

$$\sum_i K_i^\dagger K_i \leq \mathbb{1}. \quad (12)$$

If  $\sum_i K_i^\dagger K_i = \mathbb{1}$ , the operation is called a *nonselective operation*; if  $\sum_i K_i^\dagger K_i < \mathbb{1}$ , it is *selective*. Such a representation of a completely positive mapping of the state space into itself is called a *Kraus representation* of the mapping; and the operators  $\{K_i\}$ , *Kraus operators*.

Selective operations reduce the trace-norm of the density operator, but this is not an issue, as normalization is only a convention. With an unnormalized density operator, we compute the expectation value of an observable represented by an operator  $A$  via

$$\langle A \rangle_\rho = \text{Tr}(\rho A) / \text{Tr}(\rho). \quad (13)$$

In the simplest case, the set of Kraus operators is a singleton; then, if the operation is nonselective, it is unitary, and we have ordinary unitary evolution. However, one can also consider stochastic processes. Suppose the state vector undergoes a stochastic transition of the following form: for some  $i$ ,

$$|\psi\rangle \rightarrow K_i |\psi\rangle, \quad (14)$$

with the probability for which transition it undergoes given by

$$p_i = \|K_i |\psi\rangle\|^2 / \|\psi\|^2. \quad (15)$$

Since these probabilities must sum to one for every vector  $|\psi\rangle$ , we must have  $\sum_i K_i^\dagger K_i = \mathbb{1}$ . For each  $i$ , we have a selective operation; one of these yields the actual state; the transition to the mixture of these candidate states, which corresponds to the proposition that someone of these transitions has occurred, without specification of which, is given by the nonselective operation obtained by summing these. It is easy to show (see Appendix) that a nonselective operation leaves the probability distribution for results of a measurement corresponding to an operator  $B$  unchanged in all states if and only if  $B$  commutes with every  $K_i$ .

This is the general form of a dynamical collapse theory in a Galilean spacetime. For any time interval  $(t, t')$ , with  $t' > t$ , the theory specifies an

operator-valued random variable, which takes on values from a set  $\{K_i\}$  (possibly with continuous index), with

$$\sum_i K_i^\dagger K_i = \mathbb{1}. \quad (16)$$

The evolution from a state at time  $t$  to a state at time  $t'$  is a stochastic process that consists of the choice of one  $K_i$ , with appropriate probabilities, and applying the corresponding operation. Unitary evolution is the deterministic special case of this schema.

For a measurement involving apparatus that couple to observables in  $A$ , the corresponding operation will be represented by Kraus operators in  $\mathcal{A}$ , and similarly for  $B$ . This permits us to distinguish between local state evolutions and nonlocal state evolutions, along the lines of the distinction, common in quantum information theory, between local operations and nonlocal operations. A change of state local to  $A$ , either due to the internal dynamics of the system or to interaction with an external influence local to  $A$ , will be represented by Kraus operators in  $\mathcal{A}$ , and similarly, for changes local to  $B$ . Changes that cannot be decomposed into local changes (such as would result from terms in the Hamiltonian that couple elements of  $\mathcal{A}$  to elements of  $\mathcal{B}$ ) count as nonlocal evolution. The condition of local action is the condition that all evolutions be composed of local evolutions. This is a condition that is satisfied by our usual relativistic quantum field theories, and by relativistic versions of dynamical collapse theories.

## 6.2 Quantum state evolution in a relativistic spacetime

This generalizes naturally to relativistic spacetimes. In Galilean spacetime, if we want to talk about state evolution, we use the Schrödinger picture, and associate with any spacelike hypersurface a state, which, in a collapse theory, yields probabilities for events to the future of the hypersurface, conditional on events to the past. In a relativistic context the analogue is what might be called the Tomonaga-Schwinger picture: with any spacelike hypersurface  $\sigma$  (which might or might not be a maximal hypersurface, as we could also consider spatially limited regions), we associate a quantum state  $\rho_\sigma$ , which yields probabilities for events in the  $D^+(\sigma)$ , the forward domain of dependence of  $\sigma$ , conditional on events in its past light cone.<sup>5</sup> We impose the

<sup>5</sup> This is the natural and most straightforward extension of the nonrelativistic Schrödinger evolution to a relativistic spacetime. It was suggested, though not advocated, by Aharonov and Albert (1984), and has been advocated by, among others, Fleming (1986, 1989, 1996), Ghirardi and Pearle (1991); Ghirardi and Grassi (1994); Ghirardi et al. (1995), Ghirardi (1999, 2000), and Myrvold (2002, 2003).

condition that operators implementing the evolution from a hypersurface  $\sigma$  to another hypersurface  $\sigma'$ , lying within  $\sigma$ 's forward domain of dependence, be implemented by operators that commute with all operators representing observables spacelike separated from the region between the two hypersurfaces.

Take two hypersurfaces  $\sigma, \sigma'$ , with  $\sigma'$  in the forward domain of dependence of  $\sigma$ , that share a common part  $\alpha$  (see Figure 1). Let  $\rho_{\alpha|\sigma}$  and  $\rho_{\alpha|\sigma'}$  be the reduced states of  $\alpha$  obtained from these states, that is, the restriction of these states to observables in  $D^+(\alpha)$ . If the evolution from  $\sigma$  to  $\sigma'$  is deterministic and unitary, then these reduced states coincide, because the unitary operator that implements the transition from  $\rho_\sigma$  to  $\rho_{\sigma'}$  commutes with all observables in  $\alpha$ 's forward domain of dependence. Suppose, however, there is a collapse between  $\sigma$  and  $\sigma'$ . For some  $i$ ,

$$\rho_{\sigma'} = K_i \rho_\sigma K_i^\dagger. \quad (17)$$

This is the state that would be used by someone who does not know the particular way in which the evolution from  $\sigma$  to  $\sigma'$  turns out to be realized, only that it is given by *some*  $K_i$ , with the appropriate probabilities. Then  $\rho_{\alpha|\sigma}$  and  $\bar{\rho}_{\alpha|\sigma'}$  coincide. But, in general, the reduced state  $\rho_{\alpha|\sigma'}$  obtained by restricting (17) to  $\alpha$  will not coincide with  $\rho_{\alpha|\sigma}$ . Let  $\bar{\rho}_{\sigma'}$  be the mixed state

$$\bar{\rho}_{\sigma'} = \sum_i K_i \rho_\sigma K_i^\dagger. \quad (18)$$

If we think that the reduced states  $\rho_{\alpha|\sigma}$  and  $\rho_{\alpha|\sigma'}$  are intended to be intrinsic states of the spacetime region  $\alpha$ , then it may seem that we have conflicting state ascriptions. And, if local beables are to be defined in terms of the state, then it might seem that there are differing accounts of the local beables ascribed to  $\alpha$ , depending on whether we take them from a state on  $\sigma$  or a state on  $\sigma'$ .<sup>6</sup> For this reason, it is important to stress:  $\rho_{\alpha|\sigma'}$  and  $\rho_{\alpha|\sigma}$  are not local beables, nor are they intrinsic states of the region  $\alpha$ ; the intrinsic state of  $\alpha$  is the state conditional on collapses in the past light cone of  $\alpha$ .<sup>7</sup>

<sup>6</sup> Something like this seems to be going on in Esfeld and Gisin (2014), who characterize this sort of view as one on which “what there is in nature depends on the choice of a hypersurface—so that different facts exist in nature relative to the choice of a particular foliation of space-time” (p. 258).

<sup>7</sup> This aspect of the view—that, in addition to the various reduced states that might be attributed to a bounded region  $\alpha$  as reduced states of larger hypersurfaces, there is also its intrinsic state, conditional on its past light cone—is missing from the discussion of Wallace and Timpson (2010, §7). This is a crucial point, because it enters into the identification of local beables for collapse theories; see next section.

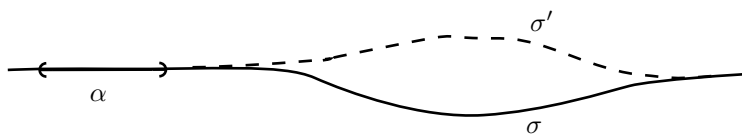


Figure 1

### 7 Local beables for relativistic collapse theories

Suppose that we had a dynamical collapse theory that produced, within a finite time, collapses to eigenstates of observables pertaining to bounded spacetime regions (perhaps a local particle number density from which one could define a local mass density). This is, after all, what the naive collapse postulate would lead one to expect. Then we could have a simple criterion for attributing local beables to systems. For a system located in a spacelike region  $\alpha$ , we would consider the state conditional on all collapses in the past light cone of  $\alpha$ , and attribute a property to the system if and only if that state is an eigenstate of the corresponding observable. That is, we would say that a system has the property  $A = a$  if and only if the state on *every* spacelike hypersurface containing  $\alpha$  is an  $A = a$  eigenstate. These would be the local beables of the theory, and our goal would be to construct a collapse theory in such a way that the totality of local beables yields a sensible world.

But collapse theories only approximate the naive collapse postulate; instead of collapsing to eigenstates of macroscopic observables, at best they yield approximations to such eigenstates. In the relativistic context, there's a matter of principle involved; it is a consequence of the Reeh-Schlieder theorem that no state of bounded energy is an eigenstate of any observable pertaining to a bounded spacetime region.

The fact that we can't expect a collapse theory to produce eigenstates of local observables led Ghirardi et al. (1990) to propose, as a criterion for attribution of local objective properties to systems, one based on approximations to eigenstates.<sup>8</sup>

We think that the appropriate attitude is the following: when considering a local observable  $A$  with its associated support we say that an individual system has the objective property  $a$  ( $a$  being an eigenvalue of  $A$ ), only when the mean value of  $P_a$  is extremely close to one, when evaluated on all spacelike hypersurfaces containing the support of  $A$  (p. 1310).

<sup>8</sup> See also Ghirardi and Pearle (1991, p. 45), Ghirardi et al. (1993, p. 362), Ghirardi and Grassi (1994, p. 417).

Ghirardi (1999, 2000) formulates this criterion in terms of the state on the past light cone.

This still leaves it open *what* observable it is such that the reduction process will lead to approximate definiteness of that observable. In the original GRW theory, it is a sort of smeared position. This does not lend itself well to a relativistic extension. Beginning in the mid-1990s, Ghirardi and collaborators have favoured mass density as the preferred variable. One defines mass density operators  $M(x)$ ; in state  $|\psi\rangle$ , the expectation value, or mean mass density, is given by

$$\mathcal{M}(x) = \langle \psi | M(x) | \psi \rangle. \quad (19)$$

Applying the criterion that an approximation to an eigenstate is good enough for property attribution, Ghirardi et al. (1995) propose that we take the mass density to be “objective” if its variance is small; in subsequent works it is said to be “accessible” if the variance is sufficiently small (Ghirardi and Grassi, 1996; Ghirardi, 1997). The idea is that a mass density with small variance is an acceptable stand-in for a classical mass density, which, of course, always has a definite value.

Combining the choice of mass density with the past light cone criterion for property attribution, the mass density that in a relativistic context is taken to be a local beable in a region  $\alpha$  is the mass density defined by the state on the past light cone of  $\alpha$  (Tumulka, 2007; Bedingham et al., 2014).

To get a sense of how the past light cone mass density functions as a local beable, consider the example raised by Einstein at the 1927 Solvay conference (Bacciagaluppi and Valentini, 2009, pp. 440–442). An electron passes through a small opening in a screen, on the other side of which is a hemispherical photographic film. The wavefunction of the electron takes the form of a spherical wave emanating from the opening, so that, when it reaches the film, the amplitude, and, consequently, the mass density, is approximately uniform on the hemisphere. When it reaches the film, the wavefunction of the electron begins to become entangled with the macroscopic degrees of freedom of the film, at which time a collapse is likely to occur, and the electron mass density becomes concentrated at some small region of the film. The past light cone mass density at another point  $p$  of the film remains the original, uniform density until the collapse event is in the past light cone of  $p$ , after which it is near-zero. There is a spherical region of near-zero mass density, centered on the position at which the collapse occurs, that spreads out from the collapse center at light speed.

On a collapse theory endowed with the past light cone mass density ontology, the mass density is a local beable, but we should not forget that

this does not exhaust the ontology; there is also the quantum state, which is not a local beable. A quantum state corresponding to a single particle is entangled across spacelike separated regions, and this entanglement shows up in anticorrelation of results; there is zero probability of collapse at two distinct parts of the screen.

Concerning the ontological status of the mass density, there are two attitudes that one can take. One—and this seems to have been the attitude behind the proposal that it be taken as “objective” or “accessible” when its variance is small—is that we still have a quantum state monist ontology. The quantum state is all that there is, and the criterion has the status of a correspondence rule or meaning postulate that tells us how to understand the mathematical apparatus of quantum theory as representing a physical world. The other, which is Ghirardi’s current view, is that “it represents an additional element which need to be posited in order to have a complete and consistent description of the world” (Ghirardi, 2007, p. 2907). On this view, collapse theories are more like the Bohm theory than might appear; on this view, a collapse theory with a mass density posits ontology above and beyond the quantum state, and the role of the quantum state is to provide dynamics for the “primitive” ontology (see Allori et al. 2008).<sup>9</sup>

## 8 A comment on Everettian theories

The mass density ontology, with approximation to an eigenstate taken as a good enough stand-in for a classical mass density, serves also for interpretations in the Oxford Everettian vein (Saunders et al., 2010; Wallace, 2012). Such interpretations suppose that there is no collapse, and all quantum state evolution is unitary. Suppose an experiment has taken place, and that the experimental apparatus has interacted with an environment; photons have been reflected from the pointers and flown out the window, or something of the sort. Then the reduced state of the laboratory takes the form of a mixture whose components are such that each component yields a mass density corresponding to approximately well-localized macroscopic objects.

In an EPR experiment, if we consider the post-decoherence state of the joint system consisting of the laboratories at *A* and *B*, we find that this takes the form of a similar mixture, with correlations between the macroscopic experimental records in the two laboratories. These correlations emerge because of the nonseparability of the quantum state from which we have de-

<sup>9</sup> My own view is that abandoning quantum state monism is premature, but that is a discussion that will have to be reserved for a later date.

rived the mass density, and as such are not locally explicable, in the sense of being mixtures over deeper descriptions on which there are no correlations.

Thus, on Everettian theories also, distant quantum correlations represent a departure from the classical expectation that all correlations be locally explicable. But, as in collapse theories, there is nothing that corresponds to action at a distance, no sense in which a choice of an instruments setting or any other variable on which one could intervene has an effect at a distance. We have some sort of nonlocality, but not action at a distance.

## 9 Conclusion

As many have pointed out, there is more than one way for a nonlocal theory to be nonlocal. I have defended the view that the difference makes a difference; for appropriate set-ups, Parameter Dependence involves a departure from relativistic causal structure, whereas theories, such as dynamical collapse theories, that satisfy Parameter Independence and exhibit only Outcome Dependence, can satisfy the requirement that Bell hoped for, namely, compatibility with relativistic causal structure at a truly fundamental level.

## 10 Acknowledgments

Work on this chapter was supported by a grant from the Foundational Questions Institute (FQXi). It was completed while the author was a Visiting Fellow at the Pittsburgh Center for the Philosophy of Science.

## 11 Appendix

### 11.1 Nonselective operations and commutation

It is easy to show that, *if* an operator  $B$  commutes with all elements  $\{K_i\}$  of a nonselective operation, then the operation leaves the statistics of  $B$ -experiments unchanged. The converse is a little (but only a little) trickier, but, since I have found that the most straightforward proof, due to Arias et al. (2002), is not as widely known as it should be, I reproduce it here.

**Theorem 1** Consider a nonselective state transition

$$\rho \rightarrow \rho' = \sum_i K_i \rho K_i^\dagger, \quad (20)$$

where

$$\sum_i K_i^\dagger K_i = \mathbb{1}. \quad (21)$$



If, for initial states  $\rho$ , the state transition (20) leaves the probabilities associated with outcomes of experiments of an observable represented by an operator  $B$  unchanged, then  $B$  commutes with each  $K_i$ .

*Proof* Suppose that, for all  $\rho$ , the state transition (20) leaves the probabilities of experiments of an observable represented by an operator  $B$  unchanged. Then it must leave the expectation value of  $B$  unchanged, as well as the variance  $(\Delta B)^2$ , and so we must have

$$\begin{aligned}\mathrm{Tr}[\rho B] &= \mathrm{Tr}[\rho' B] \\ \mathrm{Tr}[\rho B^2] &= \mathrm{Tr}[\rho' B^2].\end{aligned}\tag{22}$$

This means that

$$\mathrm{Tr}[\rho B] = \mathrm{Tr}\left[\sum_i K_i \rho K_i^\dagger B\right] = \mathrm{Tr}\left[\rho \sum_i K_i^\dagger B K_i\right].\tag{23}$$

Since this is to hold for all density operators  $\rho$ , we must have

$$\sum_i K_i^\dagger B K_i = B.\tag{24}$$

Similarly, we must have

$$\sum_i K_i^\dagger B^2 K_i = B^2.\tag{25}$$

Now consider

$$\begin{aligned}\sum_i [K_i, B]^\dagger [K_i, B] \\ = \sum_i \left( B K_i^\dagger K_i B - B K_i^\dagger B K_i - K_i^\dagger B K_i B + K_i^\dagger B^2 K_i \right)\end{aligned}\tag{26}$$

Using (21), (24), and (25), we get

$$\sum_i [K_i, B]^\dagger [K_i, B] = 0.\tag{27}$$

Since each term in (27) is a positive operator, in order to sum to zero each term must be zero, and so we must have, for each  $i$ ,

$$[K_i, B]^\dagger [K_i, B] = 0,\tag{28}$$

from which it follows that

$$[K_i, B] = 0.\tag{29}$$

□

### 11.2 Compatibility of state assignments

In this section, we will use the notion  $\rho(A)$  for the expectation value assigned to operator  $A$  by the quantum state  $\rho$ . We define a relation,  $\preceq$ , between states, by

**Definition 1**  $\rho \preceq \rho'$  if and only if, for all effects  $E$ , if  $\rho(E) = 1$  then  $\rho'(E) = 1$ .

Equivalently,

**Definition 2**  $\rho \preceq \rho'$  if and only if, for all effects  $E$ , if  $\rho'(E) > 0$  then  $\rho(E) > 0$ .

Intuitively, if  $\rho \preceq \rho'$ , anything that can happen, according to  $\rho'$ , is possible according to  $\rho$ . This relation obtains, for example, between a mixture and its components: if  $\rho$  is mixture of  $\rho'$  and some other states, then  $\rho \preceq \rho'$ .

We will say that two states are *compatible* if and only if there is no outcome of any possible experiment that is assigned probability 1 by one state and 0 by another.

**Definition 3**  $\rho \sim_c \rho'$  if and only if there is no effect  $E$  such that  $\rho(E) = 1$  and  $\rho'(E) = 0$ .

It is easy to show that this is an equivalence relation.

We also define the relation between spacelike hypersurfaces

**Definition 4**  $\sigma \leq \sigma'$  if and only if  $\sigma'$  is nowhere to the past of  $\sigma$ .

The basic assumption about the dynamics will be the *local evolution condition*:

Let  $\sigma, \sigma'$  be two spacelike Cauchy surfaces, with  $\sigma \leq \sigma'$ . Let  $\Gamma$  be the region between  $\sigma$  and  $\sigma'$ . Then there exists a family of operators  $\{K_x\}$ ,<sup>10</sup> such that

$$\int dx K_x^\dagger K_x = \mathbf{1},$$

with the following properties:

1. For any operator  $A$  representing an observable at spacelike separation from  $\Gamma$ ,  $[K_x, A] = 0$  for all  $x$ .

<sup>10</sup> This is written as if indexed by a real variable  $x$ . But the family could be points in a higher-dimensional space; take this as a stand-in for the condition that we have a measure space  $(\Omega, \mathcal{A}, \mu)$ , and an operator-valued random variable  $K$ , satisfying with the condition that

$$\int_{\Omega} d\mu(x) K_x^\dagger K_x = \mathbf{1}.$$

2. For some  $x$ , the state on  $\sigma'$  is  $\rho_x$ , defined by

$$\rho_x(A) = \rho(K_x^\dagger A K_x) / \rho(K_x^\dagger K_x).$$

3. The probability distribution over possible states on  $\sigma'$  is given by

$$\Pr(x \in \Delta) = \rho\left(\int_{\Delta} dx K_x^\dagger K_x\right).$$

We will make use of the following lemma.

**Lemma 1** For any state  $\rho$  and any effect  $E$ , if  $\rho(E) = 1$  then, for any operator  $B$ ,  $\rho(EB) = \rho(B)$ .

*Proof* We use the Cauchy-Schwartz inequality,

$$|\rho(A^\dagger B)|^2 \leq \rho(A^\dagger A) \rho(B^\dagger B). \quad (30)$$

For any effect  $F$ , if  $\rho(F) = 0$ , then, because  $F^2 \leq F$ ,  $\rho(F^2) = 0$ , and

$$|\rho(FB)|^2 \leq \rho(F^2) \rho(B^\dagger B), \quad (31)$$

and so  $\rho(FB) = 0$ . Now suppose that  $\rho(E) = 1$ . Then  $\rho(\mathbf{1} - E) = 0$ , and so, for any  $B$ ,

$$\rho((\mathbf{1} - E)B) = 0, \quad (32)$$

from which it follows that

$$\rho(EB) = \rho(B). \quad (33)$$

□

**Theorem 2** Let  $\alpha$  be an open subset of two spacelike hypersurfaces  $\sigma, \sigma'$ , with  $\sigma \leq \sigma'$ . Let  $D^+(\alpha)$  be the forward domain of dependence of  $\alpha$ , and let  $\mathcal{A}(\alpha)$  be the algebra associated with  $D^+(\alpha)$ , whose self-adjoint elements represent observables in  $D^+(\alpha)$ . Let  $\rho_{\alpha|\sigma}, \rho_{\alpha|\sigma'}$  be the restrictions of  $\rho_\sigma$  and  $\rho_{\sigma'}$  to  $\mathcal{A}(\alpha)$ . Then  $\rho_{\alpha|\sigma} \preceq \rho_{\alpha|\sigma'}$ .

*Proof* If  $\sigma$  and  $\sigma'$  are two spacelike Cauchy surfaces that share  $\alpha$ , the region between  $\sigma$  and  $\sigma'$  is spacelike separated from  $D^+(\alpha)$ . The state evolution from  $\sigma$  to  $\sigma'$  will be implemented by some operator  $K_x$ ,

$$\rho_{\sigma'}(A) = \rho_\sigma(K_x^\dagger A K_x) / \rho(K_x^\dagger K_x), \quad (34)$$

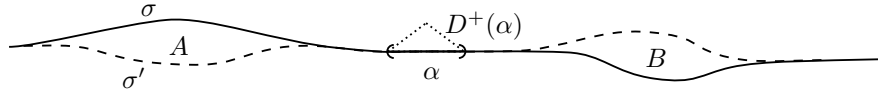
where  $K_x$  commutes with all operators representing observables in  $D^+(\alpha)$ . Thus, for any effect  $E \in \mathcal{A}(\alpha)$ ,

$$\rho_{\sigma'}(E) = \rho_\sigma(E K_x^\dagger K_x) / \rho(K_x^\dagger K_x), \quad (35)$$

If  $\rho_\sigma(E) = 1$ , by Lemma 1,  $\rho_\sigma(E K_x^\dagger K_x) = \rho_\sigma(K_x^\dagger K_x)$ , and  $\rho_{\sigma'}(E) = 1$ . □

The next theorem shows that, if  $\sigma$  and  $\sigma'$  are *any* two hypersurfaces with  $\alpha$  in common,  $\rho_\sigma$  and  $\rho_{\sigma'}$  will be compatible, in the sense of Definition 3.

Figure 2



**Theorem 3** Let  $\alpha$  be an open subset of two spacelike hypersurfaces  $\sigma, \sigma'$ . Let  $\rho_{\alpha|\sigma}, \rho_{\alpha|\sigma'}$  be the restrictions of  $\rho_\sigma$  and  $\rho_{\sigma'}$  to  $\mathcal{A}(\alpha)$ . Then  $\rho_{\alpha|\sigma} \sim_c \rho_{\alpha|\sigma'}$ .

*Proof* Let  $\sigma^+$  be a spacelike Cauchy surface containing  $\alpha$ , such that  $\sigma \leq \sigma^+$  and  $\sigma' \leq \sigma^+$ . This could be constructed, for example, by the least upper bound (in the  $\leq$  ordering) of  $\sigma$  and  $\sigma'$ . That is, it consists of the points shared by  $\sigma$  and  $\sigma'$ , plus the parts of  $\sigma$  that are to the future of  $\sigma'$ , and the parts of  $\sigma'$  that are to the future of  $\sigma$ .

$$\sigma^+ = \{x \in \sigma \mid \sigma' \leq x\} \cup \{x \in \sigma' \mid \sigma \leq x\}. \quad (36)$$

Then the evolution from  $\sigma$  to  $\sigma^+$  is through regions spacelike separated from  $\alpha$ , and hence is implemented by operators that commute with all observables in  $\mathcal{A}(\alpha)$ , and so, by Theorem 2,  $\rho_{\alpha|\sigma} \preceq \rho_{\alpha|\sigma^+}$ . By the same token,  $\rho_{\alpha|\sigma'} \preceq \rho_{\alpha|\sigma^+}$ . Therefore, if there were an effect  $E$  such that  $\rho_{\alpha|\sigma}(E) = 1$  but  $\rho_{\alpha|\sigma}(E) = 0$ , this would place contradictory demands on  $\rho_{\sigma^+}$ .  $\square$

## References

- Aharonov, Yakir, and Albert, David. 1984. Is the usual notion of time evolution adequate for quantum-mechanical systems? II. Relativistic considerations. *Physical Review D*, **29**, 228–234.
- Allori, Valia, Goldstein, Sheldon, Tumulka, Roderich, and Zanghì, Nino. 2008. On the Common Structure of Bohmian Mechanics and the Ghirardi-Rimini-Weber Theory. *The British Journal for the Philosophy of Science*, **59**, 353–389.
- Arias, A., Gheondea, A., and Gudder, S. 2002. Fixed Points of Quantum Operations. *Journal of Mathematical Physics*, **43**(12), 5872–5881.
- Bacciagaluppi, Guido, and Valentini, Antony. 2009. *Quantum Theory at the Crossroads: Reconsidering the 1927 Solway Conference*. Cambridge: Cambridge University Press.
- Bedingham, Daniel. 2011. Relativistic State Reduction Dynamics. *Foundations of Physics*, **41**, 686–704.
- Bedingham, Daniel, Dürr, Detlef, Ghirardi, GianCarlo, Goldstein, Sheldon, Tumulka, Roderich, and Zanghì, Nino. 2014. Matter Density and Relativistic Models of Wave Function Collapse. *Journal of Statistical Physics*, **154**, 623–631.
- Bell, John S. 1977. Free Variables and Local Causality. *Epistemological Letters*, **15**, 79–84. Reprinted in *Dialectica* **39** (1985), pp. 103–16, and in Bell (2004, pp. 100–104).
- Bell, John S. 1987. Are There Quantum Jumps? Pages 41–52 of: Kilmister, C.W. (ed), *Schrödinger: Centenary celebration of a polymath*. Cambridge: Cambridge University Press. Reprinted in Bell (2004, pp. 201–212).
- Bell, John S. 1990. La Nouvelle Cuisine. In: Sarlemijn, A., and Kroes, P. (eds), *Between Science and Technology*. North-Holland. Reprinted in Bell (2004, pp. 232–48).
- Bell, John S. 2004. *Speakable and Unsayable in Quantum Mechanics*. 2nd edn. Cambridge: Cambridge University Press.
- Bell, John S. 2007. The Trieste Lecture of John Stewart Bell. *Journal of Physics A: Mathematical and Theoretical*, **40**, 2919–2933.
- Dürr, Detlef, Goldstein, Sheldon, Münch-Berndl, Karin, and Zanghì, Nino. 1999. Hypersurface Bohm-Dirac Models. *Physical Review A*, **60**, 27292736.
- Dürr, Detlef, Goldstein, Sheldon, Norsen, Travis, Struyve, Ward, and Zanghì, Nino. 2014. Can Bohmian mechanics be made relativistic? *Proceedings of the Royal Society A*, **470**, 20130699.

- Esfeld, Michael, and Gisin, Nicolas. 2014. The GRW Flash Theory: A Relativistic Quantum Ontology of Matter in Space-Time? *Philosophy of Science*, **81**, 248–264.
- Evans, Peter W., Price, Huw, and Wharton, Ken B. 2013. A New Slant on the EPR-Bell Experiment. *The British Journal for the Philosophy of Science*, **634**, 297–324.
- Faye, Jan. 2010. Backward Causation. In: Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, spring 2010 edn. <http://plato.stanford.edu/archives/spr2010/entries/causation-backwards/>.
- Fleming, Gordon N. 1986. On a Lorentz Invariant Quantum Theory of Measurement. Pages 574–575 of: Greenberger, Daniel M. (ed), *New Techniques and Ideas in Quantum Measurement Theory (Annals of the New York Academy of Sciences 480)*.
- Fleming, Gordon N. 1989. Lorentz Invariant State Reduction, and Localization. Pages 112–126 of: Fine, Arthur, and Leplin, Jarrett (eds), *PSA 1988: Proceedings of the 1988 Biennial Meeting of the Philosophy of Science Association, volume two: Symposia and Invited Papers*. East Lansing, MI: Philosophy of Science Association.
- Fleming, Gordon N. 1996. Just How Radical is Hyperplane Dependence? In: Clifton, Rob (ed), *Perspectives on Quantum Reality: Non-Relativistic, Relativistic, and Field-Theoretic*. Dordrecht: Kluwer Academic Publishers.
- Ghirardi, G. C., Grassi, R., and Benatti, F. 1995. Describing the Macroscopic World: Closing the Circle within the Dynamical Reduction Program. *Foundations of Physics*, **25**, 5–38.
- Ghirardi, GianCarlo. 1997. Quantum Dynamical Reduction and Reality: Replacing Probability Densities with Densities in Real Space. *Erkenntnis*, **45**, 349–365.
- Ghirardi, GianCarlo. 1999. Some Lessons from Relativistic Reduction Models. Pages 117–152 of: Breuer, Heinz-Peter, and Petruccione, Francesco (eds), *Open Systems and Measurement in Relativistic Quantum Theory: Proceedings of the workshop held at the Istituto Italiano per gli Studi Filosofici, Naples, April 34, 1998*. Berlin: Springer.
- Ghirardi, GianCarlo. 2000. Local Measurements of Nonlocal Observables and the Relativistic Reduction Process. *Foundations of Physics*, **30**, 1337–1385.
- Ghirardi, GianCarlo. 2007. Some Reflections Inspired by My Research Activity in Quantum Mechanics. *Journal of Physics A: Mathematical and Theoretical*, **40**, 2891–2917.
- Ghirardi, GianCarlo. 2012. Does Quantum Nonlocality Irremediably Conflict with Special Relativity? *Foundations of Physics*, **40**, 1379–1395.
- Ghirardi, GianCarlo, and Grassi, Renata. 1994. Outcome Predictions and Property Attribution: the EPR Argument Reconsidered. *Studies in History and Philosophy of Science*, **25**, 397–423.
- Ghirardi, GianCarlo, and Grassi, Renata. 1996. Bohm's Theory versus Dynamical Reduction. In: Cushing, James T., Fine, Arthur, and Goldstein, Sheldon (eds), *Bohmian Mechanics and Quantum Theory: An Appraisal*. Berlin: Springer.
- Ghirardi, GianCarlo, and Pearle, Philip. 1991. Elements of Physical Reality, Nonlocality and Stochasticity in Relativistic Dynamical Reduction Models. Pages 35–47 of: Fine, Arthur, Forbes, Micky, and Wessels, Linda (eds), *PSA 1990: Proceedings of the 1990 Biennial Meeting of the Philosophy of Science Association, volume two: Symposia and Invited Papers*. East Lansing, MI: Philosophy of Science Association.

- Ghirardi, GianCarlo, Grassi, Renata, and Pearle, Philip. 1990. Relativistic Dynamical Reduction Models: General Framework and Examples. *Foundations of Physics*, **20**, 1271–1316.
- Ghirardi, GianCarlo, Grassi, Renata, Butterfield, Jeremy, and Fleming, Gordon N. 1993. Parameter Dependence and Outcome Dependence in Dynamical Models for State Vector Reduction. *Foundations of Physics*, **23**, 341–364.
- Jarrett, Jon. 1984. On the physical significance of the locality conditions in the Bell arguments. *Noûs*, **18**, 569–589.
- Myrvold, Wayne C. 2002. On Peaceful Coexistence: Is the Collapse Postulate Incompatible with Relativity? *Studies in History and Philosophy of Modern Physics*, **33**, 435–466.
- Myrvold, Wayne C. 2003. Relativistic Quantum Becoming. *The British Journal for the Philosophy of Science*, **53**, 475–500.
- Norsen, Travis. 2009. Local Causality and Completeness: Bell vs. Jarrett. *Foundations of Physics*, **39**, 273–294.
- Redhead, Michael. 1987. *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford University Press.
- Saunders, Simon, Barrett, Jonathan, Kent, Adrian, and Wallace, David (eds). 2010. *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford: Oxford University Press.
- Shimony, Abner. 1978. Metaphysical Problems in the Foundations of Quantum Mechanics. *International Philosophical Quarterly*, **18**, 3–17.
- Shimony, Abner. 1984. Controllable and Uncontrollable Non-Localities. Pages 225–30 of: S. Kamefuchi, et al. (ed), *Foundations of Quantum Mechanics in Light of New Technology*. Tokyo: The Physical Society of Japan. Reprinted in Shimony (1993, pp. 130–138).
- Shimony, Abner. 1986. Events and Processes in the Quantum World. Page 182–203 of: Penrose, R., and Isham, C. J. (eds), *Quantum Concepts in Space and Time*. Oxford University Press. Reprinted in Shimony (1993, pp. 140–162).
- Shimony, Abner. 1993. *Search for a Naturalistic Worldview, Volume II: Natural Science and Metaphysics*. Cambridge University Press.
- Shimony, Abner, Horne, Michael A., and Clauser, John F. 1976. Comment on “The Theory of Local Beables”. *Epistemological Letters*, **13**, 1–8. Reprinted in *Dialectica* **39** (1985), pp. 97–102, and in Shimony (1993, pp. 163–167).
- Skyrms, Brian. 1984. EPR: Lessons for Metaphysics. *Midwest Studies in Philosophy*, **9**, 245–255.
- Tumulka, Roderich. 2006. A Relativistic Version of the Ghirardi-Rimini-Weber Model. *Journal of Statistical Physics*, **125**, 825–844.
- Tumulka, Roderich. 2007. The unromantic pictures of quantum theory. *Journal of Physics A: Mathematical and Theoretical*, **40**, 3245–3273.
- Wallace, David. 2012. *The Emergent Multiverse*. Oxford University Press.
- Wallace, David, and Timpson, Christopher G. 2010. Quantum Mechanics on Spacetime I: Spacetime State Realism. *The British Journal for the Philosophy of Science*, **61**, 697–727.
- Woodward, James. 2013. Causation and Manipulability. In: Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, winter 2013 edn. <http://plato.stanford.edu/archives/win2013/entries/causation-mani/>.