An Exchange on Local Beables*

Summary

a) Bell tries to formulate more explicitly a notion of "local causality": correlations between physical events in different space-time regions should be explicable in terms of physical events in the overlap of the backward light cones. It is shown that ordinary relativistic quantum field theory is not locally causal in this sense, and cannot be embedded in a locally causal theory.

b) Clauser, Horne and Shimony criticize several steps in Bell's argument that any theory of local "beables" is incompatible with quantum mechanics. It is contended that the Clauser-Horne derivation of a Bell-type inequality circumvents his weak steps. The Clauser-Horne derivation must assume that there are no undetected correlations between choices of controllable variables in two space-like separated regions. Methodological considerations support this assumption.

c) In response to criticism by Shimony, Horne, and Clauser, Bell tries to clarify the argument of "The theory of local beables", and to defend as permissible the hypothesis of free variables.

d) Bell's reply to an earlier criticism by Shimony, Clauser, and Horne is answered. The convergence of Bell's position towards theirs is noted.

Résumé

a) Bell tente de formuler de façon plus explicite une notion de «causalité locale»: les corrélations entre événements physiques dans des régions spatio-temporelles différentes devraient être explicables à partir d'événements physiques dans l'intersection de leurs cônes de lumières passés. Il montre que la théorie quantique des champs relativiste ordinaire n'est pas localement causale dans ce sens et ne peut donc être intégrée dans une théorie localement causale.

b) Shimony, Horne et Clauser critiquent plusieurs pas dans la démonstration de Bell que toute théorie postulant des «beables» locales est incompatible avec la mécanique quantique. On prétend que la manière dont Clauser-Horne ont établi une inégalité du type Bell évite ces points faibles. La démonstration de Clauser-Horne doit supposer qu'il n'y a pas de corrélations non détectées entre les choix des variables contrôlables dans deux régions dont la séparation est du genre espace. Des considérations méthodologiques justifient ce présupposé.

c) En réponse aux critiques de Shimony, Horne et Clauser, Bell tente de clarifier le raisonnement de «The theory of local beables» et de défendre la légitimité de l'hypothèse de variables libres.

d) Réplique de Shimony à cette réponse de Bell. La convergence entre les positions des deux parties est soulignée.

Zusammenfassung

a) Bell versucht, in mehr expliziter Weise den Begriff der «lokalen Kausalität» zu bestimmen: die Korrelationen unter physikalischen Ereignissen in verschiedenen raum-zeitlichen Ge-

* Selected papers from *Epistemological Letters Hidden Variables and Quantum Uncertainty,* ed. by F. Bonsack, Institut de la Méthode, C.P. 1081, 2501 Bienne: No. 9 (1976), No. 13 (1976), No. 15 (1977), No. 18 (1978)

bieten sollten aufgrund von physikalischen Ereignissen in der Überschneidung ihrer vergangenen Lichtkegel erklärt werden. Er zeigt, dass die gewöhnliche relativistische Quantenfeldtheorie nicht in diesem Sinne lokal kausal ist und dass sie deshalb nicht in eine lokal kausale Theorie integriert werden kann.

b) Shimony, Horne und Clauser kritisieren mehrere Schritte in Bells Beweisgang, dass jede Theorie, die lokale «beables» postuliert, mit der Quantenmechanik unvereinbar sei. Es wird behauptet, dass die Art, wie Clauser-Horne eine Ungleichung des Bellschen Typs darlegen, diese Mängel vermeidet. Das Argument setzt voraus, dass es keine unaufgedeckten Korrelationen gibt zwischen der Wahl von in zwei raumartig getrennten Gebieten kontrollierbaren Variablen. Methodologische Überlegungen rechtfertigen diese Voraussetzung.

c) Bell antwortet auf diese Kritik, indem er den Gedankengang der früheren Arbeit klärt und die Legitimität seiner Hypothese verteidigt.

d) Shimony antwortet und unterstreicht die Konvergenz der Positionen.

J.S. Bell - *The Theory of Local Beables* **

Introduction - The Theory of Local Beables

This is a pretentious name for a theory which hardly exists otherwise, but which ought to exist. The name is deliberately modelled on "the algebra of local observables". The terminology, *be*-able as against *observ*-able, is not designed to frighten with metaphysic those dedicated to realphysic. It is chosen rather to help in making explicit some notions already implicit in, and basic to, ordinary quantum theory. For, in the words of Bohr [1], "it is decisive to recognize that, however far the phenomena transcend the scope of classical physical explanation, the account of all evidence must be expressed in classical terms". It is the ambition of the theory of local beables to bring these "classical terms" into the mathematics, and not relegate them entirely to the surrounding talk.

The concept of "observable" lends itself to very precise *mathematics* when identified with "self-adjoint operator". But physically, it is a rather woolly concept. It is not easy to identify precisely which physical processes are to be given the status of "observations" and which are to be relegated to the limbo between one observation and another. So it could be hoped that some increase in precision might be possible by concentration on the *be*ables, which can be described in "classical terms", because they are there. The beables must include the settings of switches and knobs on experimental equipment, the currents in coils, and the readings of instruments. "Observables" must be *made,* somehow, out of beables. The theory of local beables should contain, and give precise physical meaning to, the algebra of local observables.

** Presented at the sixth GIFT Seminar Jaca, June 2—7, 1975

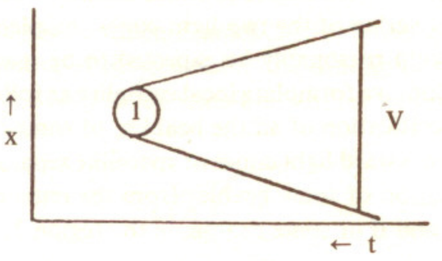
The word "beable" will also be used here to carry another distinction, that familiar already in classical theory between "physical" and "non-physical" quantities. In Maxwell's electromagnetic theory, for example, the fields $\vec{E}$ and $\vec{H}$ are "physical" (beables, we will say) but the potentials $\vec{A}$ and Φ are "non-physical". Because of gauge invariance the same physical situation can be described by very different potentials. It does not matter that in Coulomb gauge the scalar potential propagates with infinite velocity. It is not really supposed to *be* there. It is just a mathematical convenience.

One of the apparent non-localities of quantum mechanics is the instantaneous, over all space, "collapse of the wave function" on "measurement". But this does not bother us if we do not grant beable status to the wave function. We can regard it simply as a convenient but inessential mathematical device for formulating correlations between experimental procedures and experimental results, i.e., between one set of beables and another. Then its odd behaviour is as acceptable as the funny behaviour of the scalar potential of Maxwell's theory in Coulomb gauge.

We will be particularly concerned with *local* beables, those which (unlike for example the total energy) can be assigned to some bounded space time region. For example, in Maxwell's theory the beables local to a given region are just the fields $\vec{E}$ and $\vec{H}$, in that region, and all functionals thereof. It is in terms of local beables that we can hope to formulate some notion of local causality. Of course we may be obliged to develop theories in which there *are* no strictly local beables. That possibility will not be considered here.

1) *Local determinism*

In Maxwell's theory, the fields in any spacetime region 1 are determined by those in any space region V, at some time t, which fully closes the backward light cone of 1:



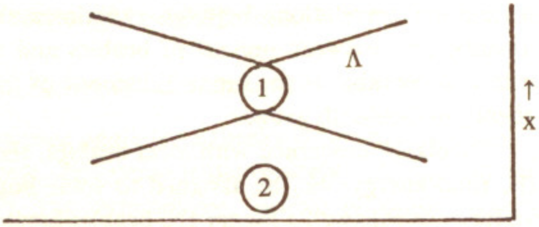
Because the region V is limited, localized, we will say the theory exhibits *local determinism*. We would like to form some notion of *local causality* in theories which are not deterministic, in which the correlations prescribed by the theory, for the beables, are weaker.

2) *Local causality*

Consider a theory in which the assignment of values to some beables Λ implies, not necessarily a particular value, but a probability distribution, for another beable A. Let

$$\{A \mid \Lambda\}$$

denote the probability of a particular value A given particular values Λ. Let A be localized in a space-time region 1. Let B be a second beable localized in a second region 2 separated from 1 in a spacelike way:



Now my intuitive notion of local causality is that events in 2 should not be "causes" of events in 1, and vice versa. But this does not mean that the two sets of events should be uncorrelated, for they could have common causes in the overlap of their backward light cones. It is perfectly intelligible then that if Λ in (1) does not contain a complete record of events in that overlap, it can be usefully supplemented by information from region 2. So in general it is expected that

$$\{A \mid \Lambda, B\} \neq \{A \mid \Lambda\} \tag{1}$$

However, in the particular case that Λ contains already a *complete* specification of beables in the overlap of the two light cones, supplementary information from region 2 could reasonably be expected to be redundant. So, with some change of notation, we formulate local causality as follows:

Let N denote a specification of *all* the beables, of some theory, belonging to the overlap of the backward light cones of spacelike separated regions 1 and 2. Let Λ be a specification of some beables from the remainder of the backward light cone of 1, and B of some beables in the region 2. Then in a *locally causal theory*

$$\{A \mid \Lambda, N, B\} = \{A \mid \Lambda, N\} \tag{2}$$

whenever both probabilities are given by the theory.

3) *Quantum mechanics is not locally causal*

Ordinary quantum mechanics, even the relativistic quantum field theory, is not locally causal in the sense of (2). Suppose, for example, we have a radioactive nucleus which can emit a single α-particle, surrounded at a considerable distance by α-particle counters. So long as it is not specified that some *other* counter registers, there is a chance for a particular counter that *it* registers. But if it is specified that some other counter does register, even in a region of space-time outside the relevant backward light cone, the chance that the given counter registers is zero. We simply do not have (2). Could it be that here we have an incomplete specification of the beables N? Not so long as we stick to the list of beables recognized in ordinary quantum mechanics — the settings of switches and knobs and currents needed to prepare the initial unstable nucleus. For these are completely summarized, in so far as they are relevant for predictions about counter registering, in so far as such predictions are possible in quantum mechanics, by the wave function.

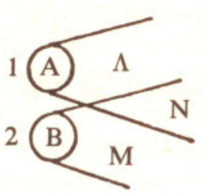
But could it not be that quantum mechanics is a fragment of a more complete theory, in which there are other ways of using the given beables, or in which there are additional beables — hitherto "hidden" beables? And could it not be that this more complete theory has local causality? Quantum mechanical predictions would then apply not to given values of all the beables, but to some probability distribution over them, in which the beables recognized as relevant by quantum mechanics are held fixed. We will investigate this question, and answer it in the negative.

4) *Locality inequality*

Consider a pair of beables A and B, belonging respectively to regions 1 and 2 with spacelike separation, which happen by definition to have the property

$$|A| \leq 1 \qquad\qquad |B| \leq 1 \tag{3}$$

Consider the situation in which beables Λ , M, N are specified, where N is a *complete* specification of the beables in the overlap of the light cones, and Λ and M belong respectively to the remainders of the two light cones



Consider the joint probability distribution

$$\{A, B \mid \Lambda, M, N\} \tag{4}$$

By a standard rule of probability, it is equal to

$$\{A \mid \Lambda, M, N, B\}\{B \mid \Lambda, M, N\} \tag{5}$$

which, by (2), is the same as

$$\{A \mid \Lambda, N\}\{B \mid M, N\} \tag{6}$$

This says simply that correlations between A and B can arise only because of common causes N.

Consider now the expectation value of the product AB

$$p(\Lambda, M, N) = \sum_{A, B} A B \{A \mid \Lambda, N\}\{B \mid M, N\} \tag{7}$$

(where the summation stands also, if necessary, for integration)

$$= \bar{A}(\Lambda, N)\,\bar{B}(M, N) \tag{8}$$

where $\bar{A}$ and $\bar{B}$ are functions of the variables indicated, and

$$|\bar{A}| \le 1 \qquad\qquad |\bar{B}| \le 1 \tag{9}$$

for all values of the arguments. Let Λ' and M' be alternative specifications, of the same regions, to ∧ and M.

$$\begin{aligned} p(\Lambda, M, N) \pm p(\Lambda, M', N) &= \bar{A}(\Lambda, N)[\bar{B}(M, N) \pm \bar{B}(M', N)] \\ p(\Lambda', M, N) \pm p(\Lambda', M', N) &= \bar{A}(\Lambda', N)[\bar{B}(M, N) \pm \bar{B}(M', N)] \end{aligned} \tag{10}$$

whence, using (9),

$$\begin{aligned} |p(\Lambda, M, N) \pm p(\Lambda, M', N)| &\le |\bar{B}(M, N) \pm \bar{B}(M', N)| \\ |p(\Lambda', M, N) \pm p(\Lambda', M', N)| &\le |\bar{B}(M, N) \pm \bar{B}(M', N)| \end{aligned} \tag{11}$$

so that finally, again invoking (9), and $|a + b| + |a - b| \le 2\,\mathrm{Max}(|a|, |b|)$,

$$|p(\Lambda, M, N) \pm p(\Lambda, M', N)| + |p(\Lambda', M, N) \mp p(\Lambda', M', N)| \le 2 \tag{12}$$

Suppose now the specifications Λ, M, N are each given in two parts

$$\begin{aligned} \Lambda &\equiv (a, \lambda) \\ M &\equiv (b, \mu) \\ N &\equiv (c, \nu) \end{aligned}$$

Consider the joint probability distribution

$$\{A , B \mid \Lambda , M , N\} \tag{4}$$

By a standard rule of probability, it is equal to

$$\{A \mid \Lambda , M , N , B\}\{B \mid \Lambda , M , N\} \tag{5}$$

which, by (2), is the same as

$$\{A \mid \Lambda , N\}\{B \mid M , N\} \tag{6}$$

This says simply that correlations between A and B can arise only because of common causes N.

Consider now the expectation value of the product AB

$$p(\Lambda, M, N) = \sum_{A, B} A B \{A \mid \Lambda , N\}\{B \mid M , N\} \tag{7}$$

(where the summation stands also, if necessary, for integration)

$$= \bar{A}(\Lambda, N)\,\bar{B}(M, N) \tag{8}$$

where $\bar{A}$ and $\bar{B}$ are functions of the variables indicated, and

$$|\bar{A}| \leq 1 \qquad\qquad |\bar{B}| \leq 1 \tag{9}$$

for all values of the arguments. Let Λ' and M' be alternative specifications, of the same regions, to $\wedge$ and M.

$$\begin{aligned} p(\Lambda, M, N) \pm p(\Lambda, M', N) &= \bar{A}(\Lambda, N)[\bar{B}(M, N) \pm \bar{B}(M', N)] \\ p(\Lambda', M, N) \pm p(\Lambda', M', N) &= \bar{A}(\Lambda', N)[\bar{B}(M, N) \pm \bar{B}(M', N)] \end{aligned} \tag{10}$$

whence, using (9),

$$\begin{aligned} |p(\Lambda, M, N) \pm p(\Lambda, M', N)| &\leq |\bar{B}(M, N) \pm \bar{B}(M', N)| \\ |p(\Lambda', M, N) \pm p(\Lambda', M', N)| &\leq |\bar{B}(M, N) \pm \bar{B}(M', N)| \end{aligned} \tag{11}$$

so that finally, again invoking (9), and $|a + b| + |a - b| \leq 2\,\mathrm{Max}(|a|, |b|)$,

$$|p(\Lambda, M, N) \pm p(\Lambda, M', N)| + |p(\Lambda', M, N) \mp p(\Lambda', M', N)| \leq 2 \tag{12}$$

Suppose now the specifications Λ, M, N are each given in two parts

$$\begin{aligned} \Lambda &\equiv (a, \lambda) \\ M &\equiv (b, \mu) \\ N &\equiv (c, \nu) \end{aligned}$$

where we are particularly interested in the dependence on a, b, c, while λ, μ, ν, are averaged over some probability distributions — which may depend on a, b, c. In the comparison with quantum mechanics, we will think of a, b, c, as variables which specify the experimental set-up in the sense of quantum mechanics, while λ, μ, ν, are in that sense either hidden or irrelevant.
Define

$$P(a, b, c) = \overline{p\big((a, \lambda), (b, \mu), (c, \nu)\big)} \tag{13}$$

where the bar denotes the averaging over (λ, μ, ν) just described. Now applying again the locality hypothesis (3), the distribution of λ and ν must be independent of b, μ — the latter being outside the relevant backward light cones.
So

$$|P(a, b, c) \pm P(a, b', C)|$$
$$\leq \overline{|p\big((a,\lambda),(b,\mu),(c,\nu)\big) \pm p\big((a,\lambda),(b',\mu'),(c,\nu)\big)|} \tag{14}$$

- because the mod of the average is less than the average of the mod. In the same way

$$|P(a', b, c) \mp P(a', b', c)|$$
$$\leq \overline{|p\big((a', \lambda'), (b, \mu), (c, \nu)\big) \mp p\big((a', \lambda'), (b', \mu'), (c, \nu)\big)|} \tag{15}$$

Finally then, from (14), (15) and (12),

$$|P(a, b, c) \mp P(a, b', c)| + |P(a', b, c) \pm P(a', b', c)| \leq 2 \tag{16}$$

5) *Quantum mechanics*

Quantum mechanics, however, gives certain correlations which *do not satisfy* the locality inequality (16).

Suppose, for example, a neutral pion is produced, by some experimental device, in some small space-time region 3. It quickly decays into a pair of photons. Suppose we have photon counters in space-time regions 1 and 2 so located with respect to 3 that when one photon falls on 1, the second falls (or nearly always does) on 2. If the π° is at rest the counters must be equally far away in opposite directions and their sensitive times appropriately delayed. Of course, both photons will often miss both counters. Suppose finally that both counters are behind filters which pass only photons with specified linear polarization, say at angles θ and ϕ respectively to some plane containing the axis joining the two counters.

Let us calculate according to quantum mechanics the probability of the various possible responses of the counters. If $| \theta \rangle$ denotes a photon linearly polarized at un angle θ, then for the photons going towards the counters the combined spin state is

$$| s \rangle = \frac{1}{\sqrt{2}} |0\rangle \, |\pi/2\rangle - \frac{1}{\sqrt{2}} |\pi/2\rangle \, |0\rangle \tag{17}$$

where first and second kets in each term refer to the photons going towards regions 1 and 2, respectively. This form is dictated by considerations of parity and angular momentum. The probability that such photons pass the filters is then proportional to

$$\begin{aligned} &\tfrac{1}{2} \left| \langle \theta | 0 \rangle \langle \phi | \pi/2 \rangle - \langle \theta | \pi/2 \rangle \langle \phi | 0 \rangle \right|^2 \\ &= \tfrac{1}{2} \left| \cos\theta \sin\phi - \sin\theta \cos\phi \right|^2 \\ &\quad = \tfrac{1}{2} \left| \sin(\theta - \phi) \right|^2 \end{aligned} \tag{18}$$

The corresponding factor for photon 1 to pass and photon 2 not is

$$\begin{aligned} &\tfrac{1}{2} \left| \langle \theta | 0 \rangle \langle \phi + \pi/2 | \pi/2 \rangle - \langle \theta | \pi/2 \rangle \langle \phi + \pi/2 | 0 \rangle \right|^2 \\ &\qquad = \tfrac{1}{2} \left| \cos(\theta - \phi) \right|^2 \end{aligned} \tag{19}$$

and so on. The probabilities for the various possible counting configurations are then

$$\begin{aligned} \varrho\,(\text{yes, yes}) &= \frac{x\Omega}{4\pi} \tfrac{1}{2} \left| \sin(\theta - \phi) \right|^2 \\ \varrho\,(\text{yes, no}) &= \frac{x\Omega}{4\pi} \tfrac{1}{2} \left| \cos(\theta - \phi) \right|^2 \\ \varrho\,(\text{no, yes}) &= \frac{x\Omega}{4\pi} \tfrac{1}{2} \left| \cos(\theta - \phi) \right|^2 \\ \varrho\,(\text{no, no}) &= \frac{x\Omega}{4\pi} \tfrac{1}{2} \left| \sin(\theta - \phi) \right|^2 + x\left(1 - \frac{\Omega}{4\pi}\right) + (1 - x) \end{aligned} \tag{20}$$

where x is the probability that the π° production mechanism actually works, Ω the (small) solid angle subtended by each counter at the production point, and no allowance has been made for bad timing, bad placing, or inefficient counting.

Now let us count $A = \pm 1$ for (yes/no) at 1 and $B = \pm 1$ for (yes/no) at 2. Then the quantum mechanical mean value of the product is

$$\begin{aligned} P(\theta, \phi) &= \varrho\,(\text{yes, yes}) + \varrho\,(\text{no, no}) - \varrho\,(\text{yes, no}) - \varrho\,(\text{no, yes}) \\ &= 1 - \frac{x\Omega}{4\pi} \left(1 + \cos 2(\theta - \phi)\right) \end{aligned} \tag{21}$$

so that

$$|P(\theta, \phi) - P(\theta, \phi')| + P(\theta', \phi) + P(\theta', \phi') - 2 =$$

$$\frac{x\Omega}{4\pi} \left\{ \left| \cos 2(\theta - \phi) - \cos 2(\theta - \phi') \right| - \cos 2(\theta' - \phi) - \cos 2(\theta' - \phi') - 2 \right\} \tag{22}$$

The right-hand side of this expression is sometimes positive. Take in particular

$$\phi = 0, 2\theta = \tfrac{\pi}{4}, - 2\phi' = \tfrac{\pi}{2}, 2\theta' = \tfrac{3\pi}{4} \tag{23}$$

in which case the factur in curly brackets is

$$\{ \qquad \} = \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}} - 2 = + 2(\sqrt{2} - 1) \tag{24}$$

But if quantum mechanics were embeddable in a locally causal theory (16) would apply, with a $\rightarrow \theta$, b $\rightarrow \phi$, and c the implicit specification of the production mechanism, held fixed in (22). The right-hand side of (22) should then be *negative*. So quantum mechanics is *not* embeddable in a locally causal theory as formulated above.

6) *Experiments*

These considerations have inspired a number of experiments. The accuracy of quantum mechanics on the atomic scale makes it hard to believe that it could be seriously wrong on that scale in some hitherto undiscovered way. The ground state of the helium atom, for example, is just the kind of correlated wave function which is embarrassing, and its energy comes out right to very high accuracy. But perhaps it is sensible to verify that theses curious correlations persist over macroscopic distances.

Experiments so far performed do not at all approach the ideal in which the settings of the instruments are determined only while the particles are in flight. When they are decided in advance, in space time regions projecting into the overlap of the backward light cones, (16) does not follow from (12). For it was supposed in (12) that the complete specification N of the overlap is the same for the various cases compared. So one can imagine a theory which is locally causal in our sense but still manages to agree with quantum mechanics for static instruments. But it would have to contain a very clever mechanism by which the result registered by one instrument depends, after a suitable time lapse, on the setting of an arbitrarily distant instrument. So static experiments are also quite interesting.

Practical experiments are far removed from the ideal in other directions also. Geometrical and other inefficiencies lead to counters registering (no, no) with overwhelming probability, (yes, yes) very seldom, and (yes, no) and (no, yes) with probabilities only weakly dependent on the settings of the instruments. Then from (21)

$$P = 1 - \epsilon^2$$

with ϵ^2 weakly dependent on the variables, so that (16) is trivially satisfied. The authors in general make some more or less ad hoc extrapolation to connect the results of the pratical with the result of the ideal experiment. It is in this sense that the entirely unauthorized "Bell's limit" sometimes plotted along with experimental points has to be understood. But such experiments also are of very high interest. For if quantum mechanics is to fail somewhere, and in the absence of a monstrous conspiracy, this should show up at some point on this side of the ideal gedanken experiment.

Several of these experiments [26 27 28] show impressive agreement with quantum mechanics, and exclude deviations as large as might be suggested by the locality inequality. Another experiment [29], very similar to one of those quoted [26], is said to be in agreement with it and yet in dramatic disagreement with quantum mechanics! And another experiment [30] disagrees significantly with the quantum prediction. Of course any such disagreement, if confirmed, is of the utmost importance, and that independently of the kind of consideration we have been making here.

7) *Messages*

Suppose that we are finally obliged to accept the existence of these correlations at long range, and the gross non-locality of nature in the sense of this analysis. Can *we* then signal faster than light? To answer this we need at least a schematic theory of what *we* can do, a fragment of a theory of human beings. Suppose we can control variables like a and b above, but not those like A and B. I do not quite know what "like" means here, but suppose that beables somehow fall into two classes, "controllables" and "uncontrollables". The latter are no use for *sending* signals, but can be used for *reception*. Suppose that to A corresponds a quantum mechanical "observable", an operator $\mathcal{A}$. Then if

$$\delta\mathcal{A} / \delta b \neq o$$

we could signal between the corresponding space time regions, using a change in b to induce a change in the expectation value of $\mathcal{A}$ or of some function of $\mathcal{A}$.

Suppose next that what we do when we change b is to change the quantum mechanical Hamiltonian $\mathcal{H}$ (say by changing some extermal field), so that

$$\delta \int dt\, \mathcal{H} = \mathcal{B}\, \delta b$$

where $\mathcal{B}$ is again an "observable" (i.e., an operator) localized in the region 2 of b. Then it is an exercise [31] in quantum mechanics to show that if in a given reference system region (2) is entirely later in time than region (1)

$$\delta\mathcal{A} / \delta b = 0$$

while if the reverse is true

$$\delta\mathcal{A} / \delta b = [\mathcal{A}, -(^{i}/\hbar)\, \mathcal{B}]$$

which is again zero (for spacelike separation) in quantum field theory by the usual local commutativity condition.

So if the ordinary quantum field theory is embedded in this way in a theory of beables, it implies that faster than light signalling is not possible. In this *human* sense relativistic quantum mechanics *is* locally causal.

8) *Reservations and acknowledgements*

Of course the assumptions leading to (16) can be challenged. Equation (22) may not embody *your* idea of local causality. You may feel that only the "human" version of the last section is sensible and may see some way to make it more precise.

The space time structure has been taken as given here. How then about gravitation?

It has been assumed that the settings of instruments are in some sense free variables — say at the whim of experimenters — or in any case not determined in the overlap of the backward light cones. Indeed without such freedom I would not know how to formulate *any* idea of local causality, even the modest human one.

This paper has been an attempt to be rather explicit and general about the notion of locality, along lines only hinted at in previous publications [Refs. 2), 4), 10), 19)]. As regards the literature on the subject, I am particularly conscious of having profited from the paper of Clauser, Horne, Holt and Shimony[3], which gave the prototype of (16), and from that of Clauser and Horne[16]. As well as a general analysis of the topic this last paper contains a valuable discussion of how best to apply the inequality in practice; I am indebted to it in particular for the point that in two-body decays (as compared with three-) the basic geometrical inefficiencies enter in (22) in a relatively harmless way. I have also profited from many discussions of the whole subject with Professor B. d'Espagnat.

J.S. Bell, CERN — Genève

REFERENCES

1) N. Bohr, in Albert Einstein, Ed. Schilpp, Tudor (1951).
2) J.S. Bell, Physics *1*, 195 (1965).
3) J.F. Clauser, R.A. Holt, M.A. Horne and A. Shimony, Phys. Rev. Letters *23*, 880 (1969).
4) J.S. Bell, in Proceedings of the International School of Physics "Enrico Fermi", Course IL, Varenna 1970 (Academic Press 1971).
5) R. Friedberg (1969, unpublished) referred to by M. Jammer 17).
6) E.P. Wigner, Am. J. Phys. *38*, 1005 (1970).
7) B. d'Espagnat, Conceptual Foundations of Quantum Mechanics, Benjamin (1971).
8) K. Popper, in Perspectives in Quantum Theory, Eds. W. Yourgrau and A. Van der Merwe, M.I.T. Press (1971).

9) H.P. Stapp, Phys. Rev. *D3,* 1303 (1971).
10) J.S. Bell, Science *177,* 880 (1972).
11) P.M. Pearle, Phys. Rev. *D2,* 1418 (1970).
12) J.H. McGuire and E.S. Fry, Phys. Rev. *D7,* 555 (1972)
13) S. Freedman and E.P. Wigner, Foundations of Physics *3,* 457 (1973).
14) F.J. Belinfante, A Survey of Hidden Variable Theories, Pergamon (1973).
15) V. Capasso, D. Fortunato and F. Selleri, Int. J. of Theor. Phys. *7,* 319 (1973).
16) J.F. Clauser and M.A. Horne, Phys. Rev. *D10,* 526 (1974).
17) M. Jammer, The Philosophy of Quantum Mechanics, Wiley (1974). See in particular references to T.D. Lee (p. 308) and R. Friedberg (pp. 244, 309, 324).
18) D. Gutkowski and G. Masotto, Nuovo Cimento *22B,* 1921 (1974).
19) J.S. Bell, in The Physicist's Conception of Nature, Ed. J. Mehra and D. Reidel (1973).
20) B. d'Espagnat, Phys. Rev. *D11,* 1424 (1975).
21) G. Corleo, D. Gutkowski and G. Masotto, Palermo preprint (1974).
22) H.P. Stapp, Berkeley preprint (1975), LBL 3685.
23) D. Bohm and B. Hiley, Birkbeck preprint (1975).
24) A. Baracca, Firenze preprint (1975).
25) A. Baracca, D.J. Bohm, R.J. Hiley and A.E.G. Stuart, Birkbeck preprint (1975).
26) L.R. Kasday, in Proceedings of the International School of Physics "Enrico Fermi", Course IL, Varenna 1970 (Academic Press 1971); L.R. Kasday, J.D. Ullman and C.S. Wu, Columbia preprint (1975).
27) S.J. Freedman and J.F. Clauser, Phys. Rev. Letters *28,* 938 (1972).
28) M. Lamehi-Rachti and W. Mittig, Saclay preprint, Colloque sur un "Demi-siècle de la Mécanique Quantique", 2-5 May, 1974.
29) G. Faraci, D. Gutkowski, S. Notarrigo and A.R. Pennisi, Lettere al Nuovo Cimento *9,* 607 (1974).
30) R. A. Holt and F.M. Pipkin, Harvard preprint (1974).
31) R.E. Peierls, Proc. Roy. Soc. *A214,* 143 (1952).

A. Shimony, M.A. Horne, J.F. Clauser – Comment on "The Theory of Local Beables"*

Dr. Bell's paper, "The Theory of Local Beables", performs a valuable service in clarifying two fundamental concepts: namely, *locality* and *physical reality*. His clarification leads him to a fundamental and highly reasonable assumption, expressed in equation (2) of Sect. 2. He then attempts in Sect. 4 to prove inequality (16) as a consequence of his equation (2). Unfortunately, we believe that his proof is not correct. A counter-example shows that (16) does not follow from (2) alone. Our objections are not given in a spirit of skepticism, since (16) does follow from other reasonable assumptions of locality and physical reality. These assumptions were discussed in an earlier paper[1] and will be reconsidered in this letter.

To illustrate the falsity of his claim we consider the following local beable situation. A person concocts a set of correlation experiment data. The data consist of four columns of numbers, indexed by event number j. Two of the columns contain the apparatus parameter settings, a_j and b_j, while the other two columns contain the experimental results, A_j and B_j. These data have been so contrived as to exhibit the correlation specified by quantum mechanics. The person sends the result columns (A_j and B_j) to an apparatus manufacturer; he sends the apparatus parameter settings to the secretaries of two physicists who will perform a correlation experiment using apparatus supplied by the manufacturer. The manufacturer preprograms the apparatus simply to display in sequence the results A_j (B_j) independently of what parameter setting is employed by physicist 1 (2). As physicist 1 (2) is about to record the result of the j^{th} event, his secretary quietly whispers in his ear the suggestion that he set his apparatus parameter to the value a_j (b_j). The experimentalists thus record preprogrammed results and parameter settings *which are* consistent with the quantum mechanical prediction. Thus, when they later compare their data, they find the resulting correlation is in violation of (16). Clearly, the violation occurs even though local beables alone were responsible for the results.

Now let us examine Bell's argument in some detail, in order to see what has gone amiss. We shall first recapitulate some of his notation. Recall that he is concerned with two beables A and B, localized respectively in space-time regions 1 and 2 which are space-like separated from each other. Denote by N the full set of beables contained in the region formed by the intersection of the backward light cones of 1 and 2. Denote by Λ the full set of beables in the

* Work done under the auspices of the U.S. Energy
[1] Research & Development contract No. W-7405-Eng-48.

remainder of the backward light cone of 1 and by M those in the remainder of the backward light cone of 2. (See Fig. 1).

Bell uses the notation $\{C \mid D\}$ to mean the probability (or probability density, in case of a continuum of values for C) given D of the value C for some variable beable. (He denotes the variable by same letter C, an ambiguity in notation which causes no confusion).

Bell's formulation of local causality, Eq. (2), is essentially the following: Let C be a variable beable localized in some space-time region, large or small. This region has a unique backward light cone; let D denote *all* the beables in this backward light cone. Then

$$\{C \mid D, E\} = \{C \mid D\}$$

holds for any beables E localized in space-time regions with a space-like separation from the region of C. We fully accept that this formulation is an unambiguous and concise statement of reasonable hypotheses about locality and reality. Furthermore, some of Bell's applications of Eq. (2) are certainly legitimate, specifically the replacements

$$\{A \mid \Lambda, M, N, B\} = \{A \mid \Lambda, N\}$$
$$\{B \mid \Lambda, M, N\} = \{B \mid M, N\}$$

when he proceeds from Eq. (5) to Eq. (6). We believe, however, that Bell is incorrect in drawing certain other consequences from Eq. (2).

He proceeds by dividing Λ into two parts a and λ; M into two parts b and μ; and N into two parts c and ν, where a,b,c, are controllable variables characterizing the experimental set-up and λ, μ, ν are the respective residual parts. He then speaks of probability distributions of λ, μ, ν and says that there may depend upon a, b, c, but unfortunately he does not explicitly state what depends on what, except for the remark: "Now applying again the locality hypothesis. . ., the distribution of λ and μ must be independent of b, — the latter being outside the relevant light cones". Let us now guess the dependences which Bell has in mind (subject, of course, to correction by him). By a standard rule of probability, we have

$$\begin{aligned} &\{\lambda, \mu, \nu \mid a, b, c\} \\ &= \{\lambda \mid \mu, \nu, a, b, c\}\{\mu \mid \nu, a, b, c\}\{\nu \mid a, b, c\}. \end{aligned}$$

We conjecture that Bell now wishes to make two separate appeals to Equation (2) to obtain the following replacements;

$$\text{(i)}\ \{\lambda \mid \mu, \nu, a, b, c\} = \{\lambda \mid \nu, a, c\}$$
$$\text{(ii)}\ \{\mu \mid \nu, a, b, c\} = \{\mu \mid \nu, b, c\}$$

Both (i) and (ii) are consequences of Eq. (2) as we have reformulated it (more precisely, of slight extensions of it). For, even though the space-time region in which λ is located extends to negative infinity in time, ν, a, c are *all* the beables other than λ itself in the region of λ and in the backward light cone of this region, and μ and b *do* refer to beables with space-like separation from the λ region. Similar reasoning holds for assertion (ii).

Our guessing, however, is not finished. Bell's derivation of (16) apparently also requires a suitable assertion concerning the distribution of ν. Indeed, we see no way of filling out the outline of his argument[2] without using the following:

$$\text{(iii)}\quad \{\nu \mid \text{a, b, c}\} = \{\nu \mid \text{c}\}.$$

It seems obvious that (iii) does not follow from Eq. (2), since the space-time regions containing a and b do not have space-like separation from the region of ν. In fact, the forward light cone of the region containing ν fills all of space-time. Could (iii) perhaps be a reasonable extension of Eq. (2)? We think not, at least not at the extreme level of generality that Bell seeks, since ν is the complete specification of the region to which it refers (minus the one factor c) and consequently a specific value for ν could hardly fail to influence the subsequent values of a and b. As a result of such influences, the probability distribution over the phase space of ν values would in general be conditional upon the values of a and b. This dependence cannot be excluded without further argument. It seems to us that (iii) could be made reasonable only if the settings of a and b are the results of some spontaneous events, such as acts of free will of the experimenters. (As Bell may have assumed tacitly in his derivation of (16) and explicitly in Sect. 8). This is a logical and metaphysical possibility, which we do not intend to exclude a priori. But since Bell's argument is intended to be general, it would not be legitimate for him to justify the assertion (iii) by relying upon a metaphysics which has not been proved and which may well be false[3].

It should be noted that in the second paragraph of Sect. 6 of his letter, Bell expresses certain reservations about the decisiveness of experiments based on inequality (16). He emphasizes that, "it was supposed in (12) that the complete specification N of the overlap is the same for the various cases compared". These reservations are very close in spirit to the reservations which we have just now expressed against Bell's derivation of inequality (16) itself; thus, in a way he has anticipated our criticism.

We do not regard the flaws in Bell's argument as fatal to the enterprise of deriving an inequality which is valid for a reasonable class of local theories. We feel that such a derivation was given in the paper by Clauser and Horne[1].

Both (i) and (ii) are consequences of Eq. (2) as we have reformulated it (more precisely, of slight extensions of it). For, even though the space-time region in which λ is located extends to negative infinity in time, ν, a, c are *all* the beables other than λ itself in the region of λ and in the backward light cone of this region, and μ and b *do* refer to beables with space-like separation from the λ region. Similar reasoning holds for assertion (ii).

Our guessing, however, is not finished. Bell's derivation of (16) apparently also requires a suitable assertion concerning the distribution of ν. Indeed, we see no way of filling out the outline of his argument[2] without using the following:

$$\text{(iii)} \quad \{\nu \mid \mathrm{a}, \mathrm{b}, \mathrm{c}\} = \{\nu \mid \mathrm{c}\}.$$

It seems obvious that (iii) does not follow from Eq. (2), since the space-time regions containing a and b do not have space-like separation from the region of ν. In fact, the forward light cone of the region containing ν fills all of space-time. Could (iii) perhaps be a reasonable extension of Eq. (2)? We think not, at least not at the extreme level of generality that Bell seeks, since ν is the complete specification of the region to which it refers (minus the one factor c) and consequently a specific value for ν could hardly fail to influence the subsequent values of a and b. As a result of such influences, the probability distribution over the phase space of ν values would in general be conditional upon the values of a and b. This dependence cannot be excluded without further argument. It seems to us that (iii) could be made reasonable only if the settings of a and b are the results of some spontaneous events, such as acts of free will of the experimenters. (As Bell may have assumed tacitly in his derivation of (16) and explicitly in Sect. 8). This is a logical and metaphysical possibility, which we do not intend to exclude a priori. But since Bell's argument is intended to be general, it would not be legitimate for him to justify the assertion (iii) by relying upon a metaphysics which has not been proved and which may well be false[3].

It should be noted that in the second paragraph of Sect. 6 of his letter, Bell expresses certain reservations about the decisiveness of experiments based on inequality (16). He emphasizes that, "it was supposed in (12) that the complete specification N of the overlap is the same for the various cases compared". These reservations are very close in spirit to the reservations which we have just now expressed against Bell's derivation of inequality (16) itself; thus, in a way he has anticipated our criticism.

We do not regard the flaws in Bell's argument as fatal to the enterprise of deriving an inequality which is valid for a reasonable class of local theories. We feel that such a derivation was given in the paper by Clauser and Horne[1].

In that paper two spatially separated analyzer-detector assemblies were considered along with a source of emissions located midway between the assemblies, with each emission presumably consisting of two objectively real components. It is to be noted that the set-up described is considerably more specific than that described by Bell, and therefore assumptions concerning it can be less general and more plausible than those needed by Bell. Let K denote the complete state of one of the emissions (denoted by λ in Ref. 1). Let p_1 (K, a) be the probability that assembly 1 registers a detection event, when K is the emission state and when an adjustable parameter of the assembly is chosen to be a. Let p_2 (K, b) be similarly defined for assembly 2. Finally, let p_{12} (K, a, b) be the probability of a joint detection event given K, a, and b. Clauser and Horne make two suppositions:[4]

(1) $p_{12}(K, a, b) = p_1(K, a)\, p_2(K, b)$,

which they justify on grounds of locality and reality similar to Bell's.

(2) The distribution ϱ (K) of the emissions is independent of the settings a and b.

Supposition (2) of Clauser and Horne plays a role in their argument analogous to Bell's assertion (iii). The central question is whether the supposition (2) is more reasonable than (iii). Our contention is that it is, though we do not pretend to offer a definitive proof nor do we think that one can be given.

It is obvious to begin with, that the assumption of Clauser and Horne is very much weaker than Bell's. In his notation, their assuption is

(2) $\{K \mid a, b, c\} = \{K \mid c\}$, which can be written out as
(2′) $\int \{K \mid \nu\} \{\nu \mid a, b, c\}\, d\nu = \int \{K \mid \nu\} \{\nu \mid c\}\, d\nu$.

But this is just an integrated form of assertion (iii), $\{\nu \mid a, b, c\} = \{\nu \mid c\}$, with $\{K \mid \nu\}$ used as a weighting function in the integration. Of course, the assertion of the equality of two integrals is a much weaker statement than the equality of two integrands. But there is yet more to be said. It is well known in the statistical mechanics of extended systems that the normal dynamics of the system as well as external perturbations tend to wash out correlations between variables which are temporally or spatially well-separated, unless there are specific mechanisms of the system for maintaining these correlations. In the present case, a, b, and K are values associated with well separated events. Moreover, there is no mechanism that one can point to which sets up a correlation between the selection of parameter a (or of b) and the occurrence of an emission having state K. Therefore, even though the left hand side of (2 ') contains a factor $\{K \mid \nu\}$ and another factor $\{\nu \mid a, b, c\}$, it is reasonable that the way in which the distribution of ν is influenced by a and b is irrelevant for the distribution of K.

Bell can, of course, reply that we do not know that the distribution of emissions K is insensitive to the values of a and b, or for that matter that there are no causal links between the act of selecting a and that of selecting b. After all, the backward light cones of those two acts do eventually overlap, and one can imagine one region which controls the decision of the two experimenters who chose a and b. We cannot deny such a possibility. But we feel that it is wrong on methodological grounds to worry seriously about it if no specific causal linkage is proposed. In any scientific experiment in which two or more variables are supposed to be randomly selected, one can always conjecture that some factor in the overlap of the backward light cones has controlled the presumably random choices. But, we maintain, skepticism of this sort will essentially dismiss all results of scientific experimentation. Unless we proceed under the assumption that hidden conspiracies of this sort do not occur, we have abondoned in advance the whole enterprise of discovering the laws of nature by experimentation[5].

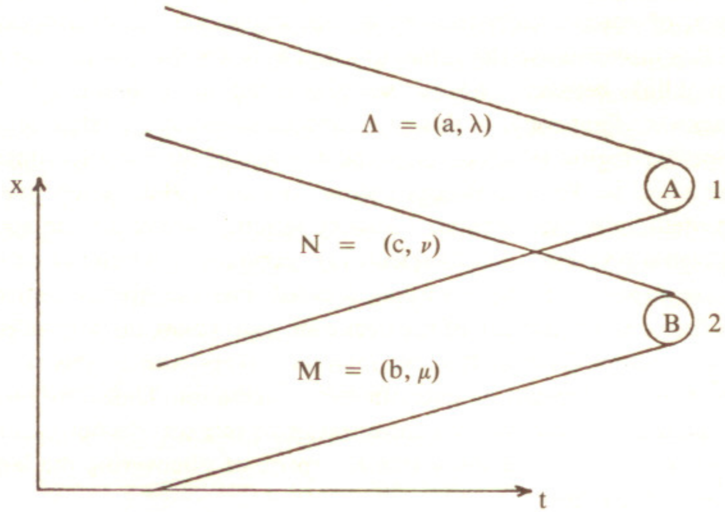
To sum up: the advantage of the Clauser-Horne approach over that of Bell's is not that it is supposition free. Rather, it is that the supposition needed is no stronger than one needs for experimental reasoning generically, and nevertheless just strong enough to yield the desired inequality.

A. Shimony, Dept of Physics & Philosophy,
Boston University

M.A. Horne, Dept of Physics,
Stonehill College,
North Easton

J.F. Clauser, Magnetic Fusion Devision,
L-386, Lawrence Livermore
Laboratory, Livermore

Fig. 1 Space-time diagram of correlation experiment beables.



REFERENCES

1. J.F. Clauser and M.A. Horne, Phys. Rev., *D10*, 526 (1974).

2. Even given (i), (ii), and (iii) there is a slip in Bell's argument for inequality (14) is incorrect, since the distribution for μ given b is different from the distribution for μ' given b'. However, given (i), (ii), and (iii), one could proceed as follows. Rewrite his Eq. 7 as

$$p(a, \lambda, b, \mu, c, \nu) = \sum_{A,B} AB\,\{A \mid a, \lambda, c, \nu\}\,\{B \mid b, \mu, c, \nu\}.$$

Multiplying both sides by $\{\lambda \mid \nu, a, c\}\,\{\mu \mid \nu, b, c\}$ and integrating we obtain

$$P(a, b, c, \nu) = \bar{A}(a, c, \nu)\,\bar{B}(b, c, \nu),$$

where

$$\bar{A}(a, c, \nu) = \int \sum_{A} A\,\{A \mid a, \lambda, c, \nu\}\,\{\lambda \mid \nu, a, c\}\, d\lambda$$

$$\bar{B}(b, c, \nu) = \int \sum_{B} B\,\{B \mid b, \mu, c, \nu\}\,\{\mu \mid \nu, b, c\}\, d\mu.$$

Then Bell's argument (9) — (16) goes through, using (iii), without any friction at step (14).

3. The objection which we have raised against assertion (iii) holds equally well against an assuption made immediately after Eq. (4.2) in an earlier paper of Bell's [in *Proceedings of the International School of Physics "Enrico Fermi", Course IL, Varenna 1970* (Academic Press 1971), p. 178] and against the form of the joint probability distribution assumed for variables λ, λ', λ'' in a paper of Shimony [In *Logic, Methodology and Philosophy of Science,* P. Suppes, et al., eds., North-Holland Publishing Company, 1973, p. 567].

4. It is interesting to note that Bell's footnote 10 in his 1970 Varenna paper (see our previous footnote for reference) can be understood as an anticipation of a proof along the lines of that of Clauser and Horne.

5. See A. Shimony, "Scientific Inference", in *The Nature and Function of Scientific Theories,* ed. R. Colodny, (U. of Pittsburgh Press, Pittsburgh, 1971).

J.S. Bell - *Free Variables and Local Causality*

It has been argued[1] that quantum mechanics is not locally causal and cannot be embedded in a locally causal theory. That conclusion depends on treating certain experimental parameters, typically the orientations of polarization filters, as free variables. Roughly speaking it is supposed that an experimenter is quite free to choose among the various possibilities offered by his equipment. But it might be that this apparent freedom is illusory. Perhaps experimental parameters and experimental results are both consequences, or partially so, of some common hidden mechanism. Then the apparent non-locality could be simulated.

This possibility is the starting point of a paper by Clauser, Horne and Shimony[2] (CHS hereafter), which is valuable in particular for a careful mathematical formulation of the assumption which excludes such a conspiracy. In this connection they severely criticize my own "theory of local beables"[1] (B hereafter). Much of their criticism is perfectly just. In B there were jumps[3] in the argument, and the assumption in question was not stated at the appropriate place, but only later and inadequately. However, I do not agree with CHS that this assumption, when carefully formulated, is an unreasonable one.

I will organize these remarks around the three phrases in which I belatedly formulated the hypothesis in B, Section 8.

1) *"It has been assumed that the settings of instruments are in some sense free variables . . ."*

For me this means that the values of such variables have implications only in their future light cones. They are in no sense a record of, and do not give information about, what has gone before. In particular they have no implications for the hidden variables ν in the overlap of the backward light cones:

$$\{\nu \mid a, b, c\} = \{\nu \mid a', b, c\} = \{\nu \mid a, b', c\} = \{\nu \mid a', b', c\} \quad (1)$$

This, as explained by CHS, is what is used in the mathematical analysis. The bracket symbol denotes the probability of particular values ν given particular values a, b, c where c lists non-hidden variables in the overlap of the backward light cones of two instruments, and a and b list non-hidden variables in the remainders of those light cones. The lists a and a' are supposed to differ in the setting of the first instrument, while b and b' are supposed to differ in the setting of the second instrument.

Note that instead of (1) CHS write, probably interpreting the symbols a little differently

$$\{\nu \mid a, b, c\} = \{\nu \mid c\}$$

With my notation, where a and b are lengthy lists of variables describing the situation outside the overlap, this would be much stronger than (1) — and not reasonable at all.

2) *"... say at the whim of experimenters ...»*

Here I would entertain the hypothesis that experimenters have free will. But according to CHS it would not be permissible for me to justify the assumption of free variables "by relying on a metaphysics which has not been proved and which may well be false". Disgrace indeed, to be caught in a metaphysical position! But it seems to me that in this matter I am just pursuing my profession of theoretical physics.

I would insist here on the distinction between analyzing various physical theories, on the one hand, and philosophising about the unique real world on the other hand. In this matter of causality it is a great inconvenience that the real world is given to us once only. We cannot know what would have happened if something had been different. We cannot repeat an experiment changing just one variable; the hands of the clock will have moved, and the moons of Jupiter. Physical theories are more amenable in this respect. We can *calculate* the consequences of changing free elements in a theory, be they only initial conditions, and so can explore the causal structure of the theory. I insist that B is primarily an analysis of certain kinds of physical theory.

A respectable class of theories, including contemporary quantum theory as it is practised, have "free" "external" variables in addition to those internal to and conditioned by the theory. These variables are typically external fields or sources. They are invoked to represent experimental conditions. They also provide a point of leverage for "free willed experimenters", if reference to such hypothetical metaphysical entities is permitted. I am inclined to pay particular attention to theories of this kind, which seem to me most simply related to our everyday way of looking at the world.

Of course there is an infamous ambiguity here, about just what and where the free elements are. The fields of Stern-Gerlach magnets could be treated as external. Or such fields and magnets could be included in the quantum mechanical system, with external agents acting only on external knobs and switches. Or the external agents could be located in the brain of the experimenter. In the latter case the sitting of the instrument is *not* itself a free variable. It is only more or less closely correlated with one, depending on how accurately the experimenter effects his intention. As he puts out his hand to the knob, his hand may shake, and may shake in a way influenced by the variables ν. Remember, however, that the disagreement between locality and quantum mechanics is large — up to a factor of $\sqrt{2}$ in a certain sense. So some hand trembling can be tolerated without much change in the conclusion. Quantification of this would require careful epsilonics.

3) *". . . or at least not determined in the overlap of the backward light cones"*

Here I must concede at once that the hypothesis becomes quite inadequate when weakened in this way. The theorem no longer follows. I was mistaken.

At this point I had in mind the possibility of exploiting the freedom, in conventional physical theories, of initial conditions. I am now embarrassed not only by the inadequacy of this particular phrase in the hypothesis, but also by the necessity of paying attention in such a study to the creation of the world[4].

Let me instead then weaken the hypothesis in a different and more practical way.

4) *". . . or at least effectively free for the purpose at hand."*

Suppose that the instruments are set at the whim, not of experimental physicists, but of mechanical random number generators. Indeed it seems less impractical to envisage experiments of this kind[5], with space-like separation between the outputs of two such devices, than to hope to realize such a situation with human operators. Could the outputs of such mechanical devices reasonably be regarded as sufficiently free for the purpose at hand? I think so.

Consider the extreme case of a "random" generator which is in fact perfectly deterministic in nature and, for simplicity, perfectly isolated. In such a device the complete final state perfectly determines the complete initial state — nothing is forgotten. And yet for many purposes, such a device is precisely a "forgetting machine". A particular output is the result of combining so many factors, of such a lengthy and complicated dynamical chain, that it is quite extraordinarily sensitive to minute variations of any one of many initial conditions. It is the familiar paradox of classical statistical mechanics that such exquisite sensitivity to initial conditions is practically equivalent to complete forgetfulness of them. To illustrate the point, suppose that the choice between two possible outputs, corresponding to a and a', depended on the oddness or evenness of the digit in the millionth decimal place of some input variable. Then fixing a or a' indeed fixes something about the input — i.e., whether the millionth digit is odd or even. But this peculiar piece of information is unlikely to be the vital piece for any distinctively different purpose, i.e., it is otherwise rather useless. With a physical shuffling machine, we are unable to perform the analysis to the point of saying just what peculiar feature of the input is remembered in the output. But we can quite reasonably assume that it is not relevant for other purposes. In this sense the output of such a device is indeed a sufficiently free variable for the purpose at hand. For this purpose the assumption (1) is then true enough, and the theorem follows.

Arguments of this kind are advanced by CHS in defending the corresponding assumption in the Clauser-Horne analysis. I do not know why they should be considered less relevant here.

Of course it might be that these reasonable ideas about physical randomizers are just wrong — for the purpose at hand. A theory may appear in which such conspiracies inevitably occur, and these conspiracies may then seem more digestible than the non-localities of other theories. When that theory is announced I will not refuse to listen, either on methodological or other grounds. But I will not myself try to make such a theory.

J.S. Bell, CERN — Genève

REFERENCES

(1) J.S. Bell, Epistemological Letters *9*, March 1976, and references therein (17.0).

(2) A. Shimony, M.A. Horne and J.F. Clauser, Epistemological Letters *13*, octobre 1976 (17.1).

(3) In particular CHS complain of a difficulty in connection with (14) and (15) in B. What is missing there is the remark that on the right-hand sides the averaging is over μ and μ', and λ and λ', separately. The operation is, explicitly,

$$\int d\lambda\, d\lambda'\, d\mu\, d\mu'\, d\nu \{\lambda \mid a, c, \nu\} \{\lambda' \mid a', c, \nu\} \{\mu \mid b, c, \nu\} \{\mu' \mid b', c, \nu\} \{\nu \mid a, b, c\}$$

According to Eq. (1) above a and/or b in the last factor may be replaced by a' and/or b' respectively. As applied for example to

$$p(\lambda, a), (\mu, b), (\nu, c)$$

which does not depend on λ' or μ', two integrations are trivial, leaving

$$\int d\lambda\, d\mu\, d\nu \{\lambda \mid a, c, \nu\} \{\mu \mid b, c, \nu \{\nu \mid a, b, c\}$$

$$\equiv \int d\lambda\, d\mu\, d\nu \{\lambda \mid a, b, \mu, c, \nu\} \{\mu \mid a, b, c, \nu\} \{\nu \mid a, b, c\}$$

(using locality)

$$\equiv \int d\lambda\, d\mu\, d\nu \{\lambda, \mu, \nu \mid a, b, c\}$$

which is the averaging involved in defining P (a, b, c).
However, I agree with CHS that an earlier style[6] averaging over λ and μ before forming the inequality, is simpler.

(4) The invocation in Ref. 1) of a *complete* account of the overlap of backward light cones is embarrassing in a related way, whether going back indefinitely or to a finite creation time — which might, by the way, have even been a creation *point*, with all backward light cones confused. R.P. Feynman in particular objected to the concept of a complete history being involved. In a more careful discussion the notion of completeness should perhaps be replaced by that of sufficient completeness for a certain accuracy, with suitable epsilonics.

(5) Important progress in this direction is being made by A. Aspect, Physical Review *D14*, 1944 (1976).

(6) J.S. Bell, in Proceedings of the International School of Physics "Enrico Fermi", Course IL, Varenna 1970.

Abner Shimony - *Reply* to Bell 17. 3*

Bell's answer[1] to criticisms[2] by Shimony, Clauser, and Horne (CHS) of his earlier paper[3] has greatly diminished the distance between their respective positions. At risk of exhibiting what Freud called "narcissism of small differences" I shall comment on a crucial passage in Bell's answer. He says:
"Consider the extreme case of a "random" generator which is in fact perfectly deterministic in nature - and, for simplicity, perfectly isolated. In such a device the complete final state perfectly determines the complete initial state - nothing is forgotten. And yet for many purposes, such a devise is precisely a "forgetting machine". A particular output is the result of combining so many factors, of such a lengthy and complicated dynamical chain, that it is quite extraordinarily sensitive to minute variations of any one of many initial conditions. It is the familiar paradox of classical statistical mechanics that such exquisite sensitivity to initial conditions is practically equivalent to complete forgetfulness of them. To illustrate the point, suppose that the choice between two possible outputs, corresponding to a and a', depended on the oddness or evenness of the digit in the millionth decimal place of some input variable. Then fixing a or a' indeed fixes something about the input - i. e., whether the millionth digit is odd or even. But this peculiar piece of information is unlikely to be the vital piece for any distinctively different purpose, i. e., it is otherwise rather useless. With a physical shuffling machine, we are unable to perform the analysis to the point of saying just what peculiar feature of the input is remembered in the output. But we can quite reasonably assume that it is not relevant for other purposes. In this sense the output of such device is indeed a sufficiently free variable for the purpose at hand. For this purpose the assumption (1) is then true enough, and the theorem follows. "

This passage is excellent up to the last sentence. The last sentence, however, seems to me to be a non-sequitur, unless the phrase "true enough" is interpreted with extreme latitude. Suppose that in Bell's idealized example of a deterministic and completely isolated generator the set of hidden variables ν is such that a would be generated and a' not. Then for all b and c

$$\{ \nu \mid a', b, c \} = 0$$

where the expressions on the left hand side denotes the probability (or probability density) that the hidden variables have the value ν , on condition that

* Work supported by the National Science Foundation

the non-hidden variables are a', b and c. On the other hand, for at least some b and c

$$\{\nu \mid a,b,c\} \neq 0.$$

Hence Bell's assumption (1), which asserts the equality of $\{\nu \mid a,b,c,\}$ and $\{\nu \mid a',b,c\}$ is false.

One can guess what Bell means by "true enough" from several sentences before the last in the passage quoted. Suppose K is some much less comprehensive feature of the generator than the complete set of hidden variables. Then from knowledge of K no inference can be drawn about whether a or a' some other output will be generated. If so, then it is reasonable to assume that

$$\{K \mid a,b,c\} = \{K \mid a',b,c\}.$$

(Rigorous examples in which this assumption is true are provided by dynamical systems having the "mixing property".[4] But this assumption is essentially the same as assumption (2) on p. 5 of Ref. 2. Thus the distance between the positions of Bell and of CHS seems to have converged to zero, but the latter can still claim to have articulated the common position with greater clarity.

There is an objection which Bell could have brought against CHS. They express a preference for the Clauser-Horne derivation[5] of Bell's inequalities over all other derivations, on grounds of generality and plausibility of assumptions. In my opinion, however, they should have recognized that Bell's derivation of 1971,[6] though very different in argument, proceeds from essentially the same assumptions. Bell assumes that the expectation value $\bar{A}$ depends only on the apparatus setting a and the hidden variables λ , and similarly for $\bar{B}$. Clauser and Horne assume that the probability p_1 of detection of emission 1 depends only upon a and λ , and similarly for the probability of detection of emission 2. When one recognizes the correct operational connections between the expectation values of observables and detection probabilities, the equivalence of these assumptions becomes clear.

A final remark concerns an entirely different point: namely, Bell's term "beable". The suffix "able" etymologically refers to a potentiality. An observable of a system is a property which can be observed, even though it may actually be the case that no one has observed it. But Bell's criterion for applying the term "beable" to things is that "they are there" (Ref. 3, p. 11); no potentiality is involved. The term "existent" would have been more accurate than his neologism. I hope that Gresham's law will not be confirmed in the present case.

A. Shimony, Boston University

REFERENCES

(1) J.S. Bell, Epistemological Letters, 15, Febr. 1977

(2) A. Shimony, M.A. Horne, and J.F. Clauser, Epistemological Letters 13, October 1976

(3) J.S. Bell, Epistemological Letters 9, March 1976

(4) See, for example, J. Lebowitz and O. Penrose, Physics Today *26,* No. 2, 23 (1973).

(5) J.F. Clauser and M.A. Horne, Phys. Rev. D10, 526 (1974)

(6) J.S. Bell, in *Foundations of Quantum Mechanics:* Proceedings of the Int. School of Physics "Enrico Fermi", Course IL, Varenna 1970, ed. by B. d'Espagnat.

SELECTIVE BIBLIOGRAPHY OF RECENT WORKS ON HIDDEN VARIABLES

I. *Surveys and Collections*

Pipkin, F.M., Advances in Atomic and Molecular Physics *14,* 281-340 (1978).

Clauser, J.F. and Shimony, A., Reports on Progress in Physics *41,* 1881-1927 (1978).

Selleri, F. and Tarozzi, G., Rivista del Nuovo Cimento *4,* no. 2, 1-53 (1981).

d'Espagnat, B., Physics Reports *110,* 202-264 (1984).

Wheeler, J.A. and Zurek, W.H., eds. *Quantum Theory and Measurement* (Princeton University Press, Princeton, N.J., 1983). Section III of this comprehensive anthology contains nine papers concerning hidden variables theories and Bell's theorem. There is an annotated bibliography.

Bell, J.S., Foundation of Physics *12,* 989-999 (1983). This is a valuable historical document, presenting the background and the heuristics of Bell's theorem in a way that is not available elsewhere.

Flato, M., Maric, Z., Milojevic, A., Sternheimer, D., and Vigier, J.-P., eds., *Quantum Mechanics, Determinism, Causality and Particles* (Reidel, Dordrecht, 1976).

II. *New Derivations of Inequalities of Bell's Type*

Garuccio, A. and Selleri, F., Foundations of Physics *10,* 209-216 (1980).

Mermin, N.D., Physical Review D *22,* 356-361 (1980).

Ogren, M., Physical Review D *27,* 1766-1773 (1983).

Garg, A. and Mermin, N.D., Foundations of Physics *14,* 1-39 (1984).

Mermin, N.D., Physics Today *38,* no. 4, 38-47 (1985).

III. *Analysis of Locality*

Fine, A., in *PSA 1980,* Asquith, A. and Giere, R., eds. (Philosophy of Science Association, E. Lansing, Mich., 1980), 535, with critical comments by A. Shimony.

Jarrett, J., Nous *18,* 569-589 (1984).

Shimony, A., in *Foundations of Quantum Mechanics in the Light of New Technology,* Kamefuchi, S. *et al,* eds. (Physical Society of Japan, Tokyo, 1984), 225-230.

Bohm, D. and Hiley, B., *ibid.,* 231-232.

IV. *Recent Experiments*

Aspect, A., Grangier, P., and Roger, G., Physical Review Letters *49,* 91-94 (1982).

Aspect, A., Dalibard, J., and Roger, G., Physical Review Letters *49,* 1804-1807 (1982). This is the most important of all the recent results concerning Bell's theorem. The authors tested a

Bell-type inequality, using an apparatus which effectively selected the orientations of polarization analyzers for each of a pair of photons while the photons were in flight.

V. *Significance of Bell's theorem and Related Experiments*

Fine, A., Physical Review Letters *48,* 291-295 (1982) and Journal of Mathematical Physics *23,* 1306-1310 (1982). These papers argue that the essential content of Bell's theorem concerns the non-existence of certain joint probability distributions, and that locality is a peripheral issue.

Shimony, A., British Journal for the Philosophy of Science *35,* 25-45 (1984). This paper contains a critique of the foregoing.

Marshall, T.W., Santos, E., and Selleri, F., Physics Letters *98A,* 5 (1983).

Piccioni, O., Bowles, P., Enscoe, C., Garland, R., and Mehlhop, W., in *Open Questions in Quantum Physics,* Tarozzi, G. and van der Merwe, A., eds. (Reidel, Dordrecht, 1984).