QUANTUM GENERALIZATIONS OF BELL'S INEQUALITY

B.S. CIREL'SON
Leningrad, U.S.S.R.

ABSTRACT. Even though quantum correlations violate Bell's inequality, they satisfy weaker inequalities of a similar type. Some particular inequalities of this kind are proved here. The more general case of instruments located in different space-time regions is also discussed in some detail.

The Einstein–Podolsky–Rosen paradox, Bell's inequality and related theoretical and experimental works (see, e.g., the survey [1]) have drawn particular attention to the correlation of two quantum observables measured by two space-like separated instruments, each one having a classical parameter (e.g., the orientation of a spin-measuring instrument). The transition probability function, i.e., the joint probability distribution of observables in some fixed state of the system, considered as a function of the above-mentioned parameters, may violate an inequality such as Bell's and, therefore, be unrealizable in classical physics or, more precisely, in all local hidden variables theories. We know today that such a transition probability function may nevertheless be realizable by suitable quantum measurements. But quantum transition probability functions obey some limitations too. The object of the present paper is to investigate these limitations in some particular cases. We show in particular that in the case of observables and parameters each having only two possible values one may limit oneself to a pair of spin one-half particles as a system and their spin components as observables. Correlations that cannot be realized in this way cannot be realized by any quantum measurements whatsoever. Besides that we attempt to treat the general case of an arbitrary set of instruments localized in certain regions of space-time, which form a partially ordered set with respect to causal dependence.

The present paper contains four theorems whose proofs will be published elsewhere and a short discussion of their physical content.

Let an observable A_k be given for each value of a parameter $k = 1,...,m$, and an observable B_l for each $l = 1,...,n$, each A_k commuting with each B_l. The following theorem characterizes all possible quantum correlations between A_k and B_l (i.e., expectation values c_{kl} of A_kB_l) in case of A_k and B_l having their spectra consisting of two points $\{-1; +1\}$ or, more generally, included in the interval $[-1; +1]$; in both cases the answer is the same and is contained in condition (4).

THEOREM 1. *The following four conditions for real numbers c_{kl}, $k = 1,...,m$, $l = 1,...,n$ are equivalent.*

(1) *There are a C^*-algebra $\mathscr{A}$ with identity, Hermitian $A_1,..., A_m, B_1,..., B_n \in \mathscr{A}$, and a state f on $\mathscr{A}$ such that, for every k,l,*

$$A_kB_l = B_lA_k; \quad -\mathbb{1} \leqslant A_k \leqslant \mathbb{1}; \quad -\mathbb{1} \leqslant B_l \leqslant \mathbb{1}; \quad f(A_kB_l) = c_{kl}\,.$$

Letters in Mathematical Physics 4 (1980) 93-100. 0377–9017/80/0042–0093 \$ 00.80.
Copyright © 1980 *by D. Reidel Publishing Company.*

(2) *There are Hermitian operators* $A_1,\ldots,A_m$, $B_1,\ldots,B_n$ *and a density matrix* W (*i.e. a positive operator with trace* 1) *in a Hilbert space* H *such that, for every* k,l,

$A_kB_l = B_lA_k$; spectrum $(A_k) \subset [-1;+1]$;

spectrum $(B_l) \subset [-1;+1]$; $\mathrm{Tr}(A_kB_lW) = c_{kl}$.

(3) *The same as* (2) *and in addition* $A_k^2 = \mathbb{1}$, $B_l^2 = \mathbb{1}$, $\mathrm{Tr}(A_kW) = 0$, $\mathrm{Tr}(B_lW) = 0$ *for every* k,l; *and* $H = H_1 \otimes H_2$, $A_k = A_k^{(1)} \otimes \mathbb{1}^{(2)}$, $B_l = \mathbb{1}^{(1)} \otimes B_l^{(2)}$, *where* $A_k^{(1)}$, $B_l^{(2)}$ *are some operators in* H_1, H_2, *respectively,* $\mathbb{1}^{(1)}$, $\mathbb{1}^{(2)}$*are identity operators; besides that all anticommutators* $A_{k_1}^{(1)} A_{k_2}^{(1)} + A_{k_2}^{(1)} A_{k_1}^{(1)}$ *and* $B_{l_1}^{(2)} B_{l_2}^{(2)} + B_{l_2}^{(2)} B_{l_1}^{(2)}$ *are scalar (i.e. proportional to* $\mathbb{1}^{(1)}$ *and* $\mathbb{1}^{(2)}$, respectively); H, H_1, H_2 *are finite dimensional, obeying*

$$2\log_2 \dim H_1 \leqslant \begin{cases} m & \text{if } m \text{ is even,} \\ m+1 & \text{if } m \text{ is odd,} \end{cases}$$

$$2\log_2 \dim H_2 \leqslant \begin{cases} n & \text{if } n \text{ is even,} \\ n+1 & \text{if } n \text{ is odd.} \end{cases}$$

(4) *There are unit vectors* $x_1,\ldots,x_m$, $y_1,\ldots,y_n$ *in a* $(m+n)$*-dimensional Euclidean space such that, for every* k,l,

$$\langle x_k, y_l \rangle = c_{kl}.$$

This theorem shows that for $m = n = 2$, the operators A_k, B_l can be chosen as 2×2 matrices obeying $A_k^2 = \mathbb{1}$, $B_l^2 = \mathbb{1}$ and having scalar anticommutators; thus, A_1 and A_2 can be interpreted as spin components along two different directions of a spin one-half particle; the same holds true for B_1, B_2 with a second particle. It is known in this case that with a suitable choice of directions and of the density matrix W one can obtain

$$c_{11} + c_{12} + c_{21} - c_{22} = \mathrm{Tr}((A_1B_1 + A_1B_2 + A_2B_1 - A_2B_2)\,W) = 2\sqrt{2},$$

whereas in the classical case all operators commute, and thus obey

$$|A_1B_1 + A_1B_2 + A_2B_1 - A_2B_2| \leqslant |B_1 + B_2| + |B_1 - B_2| \leq 2.$$

It is also known that the right-hand side $2\sqrt{2}$ is the greatest possible value for the particular linear combination of spin correlations considered above. According to Theorem 1 this implies that the inequality

$$c_{11} + c_{12} + c_{21} - c_{22} \leqslant 2\sqrt{2}$$

holds for arbitrary quantum observables A_1, A_2, B_1 and B_2 as well. There is also an elementary proof, based on a simple but lengthy calculation:

$$A_1B_1 + A_1B_2 + A_2B_1 - A_2B_2 = \frac{1}{\sqrt{2}}(A_1^2 + A_2^2 + B_1^2 + B_2^2) -$$
$$-\frac{\sqrt{2}-1}{8}((\sqrt{2}+1)(A_1 - B_1) + A_2 - B_2)^2 -$$
$$-\frac{\sqrt{2}-1}{8}((\sqrt{2}+1)(A_1 - B_2) - A_2 - B_1)^2 -$$
$$-\frac{\sqrt{2}-1}{8}((\sqrt{2}+1)(A_2 - B_1) + A_1 + B_2)^2 -$$
$$-\frac{\sqrt{2}-1}{8}((\sqrt{2}+1)(A_2 + B_2) - A_1 - B_1)^2$$
$$\leq \frac{1}{\sqrt{2}}(A_1^2 + A_2^2 + B_1^2 + B_2^2) \leq 2\sqrt{2} \cdot \mathbb{1} .$$

Theorem 1, via (4), yields comparatively simple (in general quadratic) limitations for the quantum correlations c_{kl}. This simple result is obtained only since the mean values of A_k and B_l were left undetermined. If these mean values are also considered to be given, we are able to generalize Theorem 1 only for the particular case $m = n = 2$, and even then the resulting inequalities are much more complicated. Our results are contained in the following two theorems.

THEOREM 2. *The following two definitions of a real number M, considered as a function of eight arbitrarily given real numbers $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}$, are equivalent.*

(1) $M = \sup_{A_1, A_2, B_1, B_2} (\sup \{\lambda | \lambda \in \text{spectrum}(Z)\})$

where

$$Z = \alpha_1 A_1 + \alpha_2 A_2 + \beta_1 B_1 + \beta_2 B_2 + \gamma_{11} A_1 B_1 + \gamma_{12} A_1 B_2 + \gamma_{21} A_2 B_1 + \gamma_{22} A_2 B_2 ,$$

and A_1, A_2, B_1, B_2 *are arbitrary Hermitian operators in a Hilbert space obeying*

$$A_k B_l = B_l A_k \; ; \quad \text{spectrum}(A_k) \subset [-1;+1] \; ; \quad \text{spectrum}(B_l) \subset [-1;+1]$$

for $k = 1, 2$ *and* $l = 1, 2$.

(2) $M = \inf m$,

where the infimum is taken over all $m > 0$ such that, for every complex numbers u, v obeying $|u| = 1$, $|v| = 1$, the following inequalities hold:

$$m^4 + \mu_2 m^2 + \mu_3 m + \mu_4 > 0 ,$$

$$4m^3 + 2\mu_2 m + \mu_3 > 0 ,$$

$$6m^2 + \mu_2 > 0 ,$$

where

$$\mu_2 = -(|e|^2 + |f|^2) - 2(|g|^2 + |h|^2)\,,$$

$$\mu_3 = -4\mathrm{Re}(e\bar{g}\bar{h} + f\bar{g}h)\,,$$

$$\mu_4 = |e|^2|f|^2 + (|g|^2 - |h|^2)^2 - 2\mathrm{Re}(ef\bar{g}^2 + \bar{e}fh^2)\,,$$

with

$$e = \frac{1}{2}(uv\gamma_{11} - u\bar{v}\gamma_{12} - \bar{u}v\gamma_{21} + \bar{u}\bar{v}\gamma_{22})\,,$$

$$f = \frac{1}{2}(u\bar{v}\gamma_{11} - uv\gamma_{12} - \bar{u}\bar{v}\gamma_{21} + \bar{u}v\gamma_{22})\,,$$

$$g = \frac{1}{\sqrt{2}}(u\alpha_1 - \bar{u}\alpha_2)\,,$$

$$h = \frac{1}{\sqrt{2}}(v\beta_1 - \bar{v}\beta_2)\,.$$

THEOREM 3. *The following four conditions for eight real numbers* $a_1, a_2, b_1, b_2, c_{11}, c_{12}, c_{21}, c_{22}$ *are equivalent.*

(1) *There are a* C^**-algebra* $\mathscr{A}$ *with identity, Hermitian* $A_1, A_2, B_1, B_2 \in \mathscr{A}$ *and a state* f *on* $\mathscr{A}$ *such that for* $k = 1,2$ *and* $l = 1,2$,

$$A_kB_l = B_lA_k;\ -\mathbb{1} \leq A_k \leq \mathbb{1};\ -\mathbb{1} \leq B_l \leq \mathbb{1};$$
$$f(A_k) = a_k;\ f(B_l) = b_l;\ f(A_kB_l) = c_{kl}\,.$$

(2) *There are Hermitian operators* A_1, A_2, B_1, B_2 *and a density matrix* W *in a Hilbert space* H *such that for* $k = 1,2$ *and* $l = 1,2$,

$$A_kB_l = B_lA_k;\ \text{spectrum}\,(A_k) \subset [-1;+1];$$
$$\text{spectrum}\,(B_l) \subset [-1;+1];\ \mathrm{Tr}\,(A_kW) = a_k;$$
$$\mathrm{Tr}\,(B_lW) = b_l;\ \mathrm{Tr}\,(A_kB_lW) = c_{kl}\,.$$

(3) *There are Hermitian* 2×2*-matrices* $A_1(\zeta), A_2(\zeta), B_1(\zeta), B_2(\zeta)$ *and a* 4×4 *density matrix* $W(\zeta)$*, depending in a measurable way on a parameter* $\zeta \in [0;1]$, *such that, for every* $\zeta \in [0;1]$, $k = 1,2$ *and* $l = 1,2$,

$$A_k^2(l) = \mathbb{1};\qquad B_k^2(\zeta) = \mathbb{1};\qquad A_1(\zeta)A_2(\zeta) + A_2(\zeta)A_1(\zeta)$$

and

$$B_1(\zeta)B_2(\zeta) + B_2(\zeta)B_1(\zeta)$$

are scalar; and

$$\int_0^1 \mathrm{Tr}((A_k(\zeta) \otimes \mathbb{1}_2) W(\zeta)) d\zeta = a_k ,$$

$$\int_0^1 \mathrm{Tr}((\mathbb{1}_2 \otimes B_l(\zeta)) W(\zeta)) d\zeta = b_l ,$$

$$\int_0^1 \mathrm{Tr}((A_k(A_k(\zeta) \otimes B_l(\zeta)) W(\zeta)) d\zeta = c_{kl},$$

where $\mathbb{1}_2$ *is the* 2×2 *identity matrix.*

(4) *For any eight arbitrarily given real numbers* $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}$,

$$\alpha_1 a_1 + \alpha_2 a_2 + \beta_1 b_1 + \beta_2 b_2 + \gamma_{11} c_{11} + \gamma_{12} c_{12} + \gamma_{21} c_{21} + \gamma_{22} c_{22} \leq M,$$

where M *is defined as in Theorem 2.*

We see again that it is sufficient to deal with spin one-half particles; it is necessary now, however, to use a randomization by means of a classical random parameter ζ affecting the state of the particles and the orientations of the two spin measuring instruments. Most likely our results are not generalizable for $m > 2$ or $n > 2$. In fact, any two noncommuting operators A_1, A_2, obeying $A_1^2 = \mathbb{1}, A_2^2 = \mathbb{1}$, commute with a third operator, namely with $A_1A_2 + A_2A_1$. This is of crucial importance for the proofs of Theorems 2 and 3, and has no analog for the three or more operators.

Now we turn from the case of two space-like separated (e.g., spin measuring) devices to a more general case. Of course, our results will be less concrete than those above. We shall treat an arbitrary set of local instruments following Kraus [2,3], Hellwig and Kraus [4], and Davies and Lewis [5].

Let $C_1, ..., C_n$ be regions of space-time such that for every different k,l, either any point from C_l is in the future with respect to any point from C_k (denoted by $C_k < C_l$), or vice versa ($C_l < C_k$), or any two points from C_k and C_l, respectively, are space-like separated ($C_k \sim C_l$). It is convenient to suppose that the numbering of regions conforms to their chronological order, i.e. if $k < l$ then either $C_k < C_l$ or $C_k \sim C_l$. We shall consider for each $k = 1,...,n$ a finite index set S_k and an instrument, i.e. a family $\{O_k(s_k)\}_{s_k \in S_k}$ of operations $O_k(s_k)$, such that the sum over S_k of the corresponding effects is $\mathbb{1}$. Note that each operation is a mapping $O: L_1^+(H) \to L_1^+(H)$ of the form $OW = \sum_i A_i W A_i^*$, where H is a Hilbert space describing a quantum system, $L_1^+(H)$ is the set of all positive trace-class operators, and A_i are linear operators on H obeying $F < \mathbb{1}$, where $F = \sum_i A_i^* A_i$ is the effect corresponding to O. Application of an instrument $\{O(s)\}_{s \in S}$ to the system, being in a state with a density matrix W, produces with the probability $p(s) = \mathrm{Tr}(O(s)W)$ an output $s \in S$ and a new state of the system with the density matrix $(1/p(s))\, O(s) W$. A combined examination of several local instruments requires some further conditions. We suppose that each instrument $\{O_k(s_k)\}$ is localized in the corresponding region C_k. Then the instruments must obey the following obvious condition.

Condition (A). If $C_k \sim C_l$, the instruments $\{O_k(s_k)\}$ and $\{O_l(s_l)\}$ commute, i.e.,

$$O_k(s_k)O_l(s_l)W = O_l(s_l)O_k(s_k)W$$

for every $s_k \in S_k$, $s_l \in S_l$, $W \in L_1^*(H)$.

A combined application of all instruments $\{O_k(s_k)\}$ to the density matrix W produces with the probability

$$p(\{s_k\}) = \mathrm{Tr}(O_n(s_n) \ldots O_1(s_1)W)$$

a combined output $\{s_k\}_{k=1,\ldots,n} \in \prod_k S_k$ and a new density matrix

$$\frac{1}{p(\{s_k\})} O_n(s_n) \ldots O_1(s_1)W .$$

The Condition (A) ensures that this composition of instruments is independent of their numbering (which although assumed to be chronological, is still arbitrary in case of space-like separation). But we need a stronger condition. Any operation O has a representation (see [3]) by means of an auxiliary Hilbert space H_1 describing an apparatus, a density matrix W_1 on H_1 describing the initial apparatus state, a unitary operator U_1 on $H \otimes H_1$ describing an interaction between system and apparatus, and a projection operator P on H_1 describing a property of the apparatus; we shall denote this fact as $O = Op(W_1, U_1, P)$. It is known also [3] that any instrument $\{O(s)\}_{s \in S}$ has a representation by means of a single apparatus: $O(s) = Op(W_1, U_1, P(s))$, where $\{P(s)\}_{s \in S}$ is an ideal measurement, i.e., a family of disjoint projection operators on H_1 whose sum is $\mathbb{1}$.

Condition (B). The instruments $\{O_k, (s_k)\}$ are such that there are Hilbert spaces H_k, unitary operators U_k on $H \otimes H_k$ and ideal measurements $\{P_k(s_k)\}_{s_k \in S_k}$ on H_k for $k = 1,\ldots,n$ obeying:

(B1) if $C_k \sim C_l$, then for any density matrices W_k on H_k and W_l on H_l the two instruments

$$\{Op(W_k, U_k, P_k(s_k))\}_{s_k \in S_k} \text{ and } \{Op(W_l, U_l, P_l(s_l))\}_{s_l \in S_l}$$

commute; and

(B2) for every k there is a density matrix W_k^0 on H_k such that $O_k(s_k) = Op(W_k^0, U_k, P_k(s_k))$ for all $s_k \in S_k$.

Thus, the condition (B) demands that space-like separated instruments commute not only for the prescribed initial states of apparatuses, but also for arbitrary ones. It might look surprising that a set of instruments may obey Condition (A) while violating Condition (B), but this is really a fact. Apparently, such a set of instruments is unrealizable.

Let us introduce inputs. We consider a finite parameter set R_k for each $k = 1,\ldots,n$, and suppose that each instrument depends on a corresponding parameter $r_k \in R_k$; this means that operations $O_k(r_k, s_k)$ are given for every $k = 1,\ldots,n$, $r_k \in R_k$, $s_k \in S_k$, and for each r_k, $\{O_k(r_k, s_k)\}_{s_k \in S_k}$ is an instrument. Now we define a transition probability function for fixed W:

$$p(\{s_k\} \mid \{r_k\}) = \mathrm{Tr}(O_n(r_n, s_n) \ldots O_1(r_1, s_1)W) .$$

It is convenient in the following theorem to replace the set $\{O_1,\ldots, O_n\}$ of regions by an arbitrary set, partially ordered with respect to causal dependence as above.

THEOREM 4. *Suppose that K is a partially ordered finite set, R_k and S_k are finite sets for every $k \in K$, and a real number $p(\{s_k\} \mid \{r_k\})$ is given for every $\{r_k\}_{k\in K} \in \prod_{k\in K} R_k$, $\{s_k\}_{k\in K} \in \prod_{k\in K} S_k$. Then the following two conditions are equivalent.*

(1) *There are a Hilbert space H and a family of operations $O_k(r_k, s_k)$ acting on $L_1^+(H)$ for $k \in K$, $r_k \in R_k$, $s_k \in S_k$, such that for every r_k the Condition (B) holds for $\{O_k(r_k,s_k)\}_{s_k\in S_k}$, and there is a density matrix W on H such that for every $\{r_k\}$ and $\{s_k\}$*

$$\mathrm{Tr}((\prod_{k\in k} O_k(r_k,s_k))W) = p(\{s_k\} \mid \{r_k\}) \ ,$$

where the product is conform to the order, i.e.

$$\prod_{k\in K} O_k(r_k, s_k) = O_{k_n}(r_{k_n}, s_{k_n}) \ldots O_{k_1}(r_{k_1}, s_{k_1}) \ ,$$

where $k_1,\ldots, k_n$ forms a numbering of K such that $k_i < k_j$ implies that $i < j$; the choice of this numbering does not affect the product in consequence of the Condition (B).

(2) *There are a Hilbert space H, a density matrix W and projection operators $Q_k(s_k, \{r_l\})$ on H for $k\in K$, $s_k \in S_k$, $\{r_l\}_{l\in K} \in \prod_{l\in K} R_l$, obeying*

(2a) *for every $\{r_l\}$ and k*

$$\sum_{s_k\in S_k} Q_k(s_k, \{r_l\}) = \mathbb{I} \ ,$$

summands being disjoint;

(2b) *for every $\{r_l\}$, k_1, k_2, s_{k_1}, and s_{k_2} operators $Q_{k_1}(s_{k_1}, \{r_l\})$ and $Q_{k_2}(s_{k_2}, \{r_l\})$ commute;*

(2c) *for every k and s_k the operator $Q_k(s_k, \{r_l\})$ depends in fact only on r_l having $l \leqslant k$;*

(2d) *for every $\{r_k\}$ and $\{s_k\}$*

$$\mathrm{Tr}((\prod_{k\in K} Q_k(s_k, \{r_l\}))W) = p(\{s_k\} \mid \{r_k\})$$

($\{r_l\}$ is the same as $\{r_k\}$).

The relation between Conditions (1) and (2) of Theorem 4 is somewhat similar to the relation between the interaction representation and the Heisenberg representation. Indeed, we accept in (1) that the state of the system changes because of interactions with apparatuses, whereas we do not distinguish in (2) between system and apparatuses, and accept a changeless state, but observables depending on parameters.

If one demands in addition to (2) that all $Q_k(s_k, \{r_l\})$ commute with each other even when belonging to different parameters $\{r_l\}$, then one obtains the condition of classical realizability of a transition probability function. It is easy to see that this condition is necessary in any local hidden variable theory. On the other hand, one can argue that it is also sufficient: to realize a transition probability function obeying this condition, one needs only suitable logic circuits as used in computers together with random generators. On the other hand, we have no reason to believe that in the general (noncommutative) case the necessary condition pointed out in Theorem 4 is also sufficient for physical realizability. It should be interesting to find a stronger necessary condition, as well as to look for a general sufficient condition. Perhaps a suitable candidate for this would be a 'net of abstract scatterings', i.e. an oriented graph, with Hilbert spaces corresponding to its edges and unitary operators on tensor products corresponding to its nodes. Perhaps a further development of experimental techniques connected with Bell's inequality will lead to the conclusion that any such 'net of abstract scatterings' may be approximately realized in a suitable experiment.

ACKNOWLEDGMENT

The author is grateful to L.A. Khalfin for stimulating discussions.

REFERENCES

1. Clauser, J.F., and Shimony, A., 'Bell's theorem: experimental tests and implications', *Rep. Progr. Phys.* **41**, 12, 1881–1927 (1978).
2. Kraus, K., 'General state changes in quantum theory', *Ann. Phys. (N.Y.)* **64**, 2, 311–335 (1971).
3. Kraus, K., 'Operations and effects in the Hilbert space formulation of quantum theory', *Lecture Notes in Physics,* Springer-Verlag, **29**, 206–229 (1974).
4. Hellwig, K.-E., and Kraus, K., 'Formal description of measurements in local quantum field theory', *Phys. Rev. (D)* **1**, 2, 566–571 (1970).
5. Davies, E.B., and Lewis, J.T., 'An operational approach to quantum probability', *Commun. Math. Phys.* **17**, 3, 239–260 (1970).

(Received February 20, 1980)