PRAMANA — journal of physics
© Indian Academy of Sciences
Vol. 56, Nos 2 & 3
Feb. & Mar. 2001
pp. 139–152

Bell's inequalities and Kolmogorov's axioms

DAVID ATKINSON
Institute for Theoretical Physics, University of Groningen, The Netherlands
Email: D.Atkinson@phys.rug.nl

Abstract. After recalling proofs of the Bell inequality based on the assumptions of separability and of noncontextuality, the most general noncontextual contrapositive conditional probabilities consistent with the Aspect experiment are constructed. In general these probabilities are not all positive.

Keywords. Bell; Kolmogorov; axioms.

PACS Nos 2.0; 03.65.Bz

1. Introduction

At the beginning of the 21st century, we are familiar with the idea that Euclid's axioms of geometry do not in general apply to the physical world — when a gravitational 'field' is present, Einstein's general theory of relativity has shown us how to use non-Euclidean geometry. Does quantum mechanics similarly imply that classical logic and classical probability theory also do not apply to the physical world? There is no such unanimity as in the case of geometry. Bas van Fraassen [1] states categorically:

> The new phenomena do not force violations of classical probability theory or logic.

On the other hand, Kümmerer and Maassen [2] discuss

> ... polarization experiments which show the need to extend classical probability theory.

This claim is explicitly denied by Gill [3], who takes these authors to task:

> ... though quantum reality is strange, classical probability [is] ... perfectly adequate to describe it.

In fact the dissension is not as serious as it seems. A distinction can be made between what is *required* on the one hand and what is *useful* on the other, as in the case of geometry and relativity. No departure from the axioms of Euclid is required by the fact of gravitation. It is *possible* to describe the whole content of Einstein's theory within the framework of Euclidean geometry; but it is not very *convenient* to do so, since then light does not always propagate in free space along a geodesic, and planets appear to be acted upon by 'occult' gravitational forces [4]. We shall argue that, in a similar way, it is useful to introduce nonclassical probability in the discussion of quantum mechanics, even though it is not logically necessary to do so.

After giving the axioms and definitions of classical probability theory, we shall recall [5] two proofs of the Bell inequality, one based on the requirement of separability and the other on the assumption of noncontextual counterfactual conditional probabilities. Since the Bell inequality is known to be experimentally violated [6], it follows that probabilities in nature are neither separable nor noncontextual. These features of nonseparability and contextuality are shared by quantum theory.

In previous work, we have considered nonseparability in connection with ideas of physical independence [7], this being a special case of the theory dependence of probability itself [8]. In this paper we concentrate rather on the question of contextuality. We construct the most general noncontextual conditional 'probabilities' for the Aspect experiment, going in fact beyond quantum mechanics in this respect. An explicit demonstration is provided that there are configurations in which these putative noncontextual conditional probabilities cannot all be positive.

2. Kolmogorov's axioms

The axiomatic approach to probability was formulated by Kolmogorov in 1933 in a book published in German, a Russian translation appearing three years later. We quote from the second edition of Morrison's English translation [9] verbatim:

Let E be a collection of elements $\xi,\ \eta,\ \zeta,\ \ldots$, which we shall call *elementary events,* and $\mathcal{F}$ a set of subsets of E; the elements of the set $\mathcal{F}$ will be called *random events.*

I. $\mathcal{F}$ is a field of sets.

II. $\mathcal{F}$ contains the set E.

III. To each set A in $\mathcal{F}$ is assigned a non-negative real number $P(A)$. This number $P(A)$ is called the probability of the event A.

IV. $P(E)$ equals 1.

V. If A and B have no element in common, then

$$P(A \cup B) = P(A) + P(B)\,.$$

In axiom V, we have employed the symbol $\cup$ for the union of sets, while Kolmogorov simply uses $+$; and the condition that A and B have no element in common may be expressed by $A \cap B = \emptyset$, where $\emptyset$ is the empty set, and $\cap$ is the symbol denoting intersection.

Kolmogorov adds that a system of sets is called a field if the sum, product and difference of two sets of the system also belongs to the same system. In modern notation, this means that, given $A \in \mathcal{F}$ and $B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$, $A \cap B \in \mathcal{F}$, and $A \cap B^c \in \mathcal{F}$, B^c being the complement of B with respect to E. Since $A \cap A^c = \emptyset$, it follows that $\emptyset$ belongs to $\mathcal{F}$. It is not necessary to postulate $P(\emptyset) = 0$ and $P(A) \leq 1$ for any $A \in \mathcal{F}$, for these statements are implied by the above axioms.

If E is an infinite collection of elements, then one normally restricts $\mathcal{F}$ to be such that it is closed under countable unions of sets, and one replaces axiom V by

V$'$. If A_n is a set of pairwise disjoint sets in $\mathcal{F}$, then

$$P(\cup_n A_n) = \sum_n P(A_n)\,,$$

the condition of σ-additivity.

To the above axioms are added, *as definitions*, the notions of stochastic independence and of conditional probability:

VI. The necessary *and sufficient* condition that A and B be stochastically independent events is

$$P(A \cap B) = P(A)P(B)\,.$$

Note that this is not always equivalent to *physical* independence.

VII. The conditional probability of event A, given event B, is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}\,,$$

on condition that $P(B) \neq 0$. Note that independence is quite different from disjointness, for which axiom V applies. Moreover, if A and B are independent, $P(A|B) = P(A)$.

3. Separability and Bell's inequality

Suppose that two photons are created in an angular momentum zero state, as in the experiments of Aspect *et al*. One photon falls on a polarizer at location A, behind which there is a detector, and the other photon falls on a similar polarizer at another location, B, also with a detector behind it. It is supposed that the axis of the polarizer at A is set parallel to the vector a, and that of the polarizer at B parallel to the vector b. Let $P(a)$ be the probability that the first photon is transmitted by the polarizer at A, so that it is counted by the detector. Otherwise the photon is absorbed by the polarizer and is thus not counted, the probability of this being $1 - P(a)$. Similarly, $P(b)$ and $1 - P(b)$ are the probabilities of transmission or absorption by the polarizer at B. These probabilities can be estimated by running the experiment many times and counting relative frequencies. The prediction of quantum mechanics is

$$P(a) = \tfrac{1}{2} = P(b). \tag{1}$$

Let $P(a,b)$ be the joint probability of transmission of the photons at both A and B, with polarizer settings a and b respectively. In the notation of the previous section, this would be written $P(A(a) \cap B(b))$, where $A(a)$ and $B(b)$ are the events corresponding to registering transmission at A with setting a and at B with setting b. The prediction of quantum mechanics is

$$P(a,b) = \tfrac{1}{2}\cos^2\theta\,, \tag{2}$$

where θ is the angle between the vectors a and b.

For the first derivation of the Bell inequality, it is supposed that this joint probability can be written in the form

$$P(a,b) = \int \mathrm{d}\lambda\, \rho(\lambda) P(a|\lambda) P(b|\lambda)\,,$$

which may be called the assumption of separability, with a hidden variable, λ. Here $P(a|\lambda)$ is the conditional probability density for transmission at A, given that the setting at A is a;

and the conditioning is with respect to λ. The unconditional probability for transmission at A, with setting a, can be written as

$$P(a) = \int \mathrm{d}\lambda\, \rho(\lambda) P(a|\lambda)\,, \tag{3}$$

and similarly for $P(b)$. It is required that the weight function, ρ, is non-negative and normalized:

$$\int \mathrm{d}\lambda\, \rho(\lambda) = 1 \qquad \rho(\lambda) \geq 0\,.$$

Suppose now that the experiment is repeated with new settings for the polarizers, a' and b' instead of a and b, generating new probabilities. Moreover, the combinations $\{a, b'\}$ and $\{a', b\}$ can also be realized, resulting finally in measurements of relative frequencies that estimate the joint probabilities $P(a, b)$, $P(a', b)$, $P(a, b')$ and $P(a', b')$.

We shall define the Bell coefficient, B, which involves the analogous probabilities for the four possible combinations of settings, as follows:

$$\begin{aligned} B &= P(a,b) + P(a',b) + P(a,b') - P(a',b') \\ &= \int \mathrm{d}\lambda\, \rho(\lambda) \{ P(a|\lambda)P(b|\lambda) + P(a'|\lambda)P(b|\lambda) \\ &\quad + [P(a|\lambda) - P(a'|\lambda)]\, P(b'|\lambda) \}\,. \end{aligned} \tag{4}$$

We propose to obtain an upper bound on B. If $P(a|\lambda) - P(a'|\lambda) \leq 0$, we majorize the last line of eq. (4) by omitting the term involving this difference, which is negative or zero, and we majorize $P(a|\lambda)P(b|\lambda)$ by $P(a|\lambda)$ and $P(a'|\lambda)P(b|\lambda)$ by $P(b|\lambda)$. Thus

$$B \leq \int \mathrm{d}\lambda\, \rho(\lambda) \{ P(a|\lambda) + P(b|\lambda) \}\,. \tag{5}$$

If, on the other hand, $P(a|\lambda) - P(a'|\lambda) > 0$, we majorize by replacing $P(b'|\lambda)$ by 1, which is allowed, since its coefficient in eq. (4) is in this case positive. After transposition of the resulting terms, we find

$$B \leq \int \mathrm{d}\lambda \rho(\lambda) \{ P(a|\lambda) + P(a|\lambda)P(b|\lambda) - P(a'|\lambda)[1 - P(b|\lambda)] \}\,.$$

Here the term involving $P(a'|\lambda)$ is negative or zero, and so may be omitted, and moreover we now choose to replace $P(a|\lambda)P(b|\lambda)$ by $P(b|\lambda)$. In this way we have shown that the inequality (5) is valid also in this case. Rewriting the result in terms of the unconditional probability of eq. (3), we obtain the Bell inequality in the form that we shall use it in this paper:

$$B = P(a,b) + P(a',b) + P(a,b') - P(a',b') \leq P(a) + P(b)\,. \tag{6}$$

Suppose that the settings at A and B are chosen such that the angle between a and b, between a and b' and between a' and b are all the same, say θ, while that between a' and b' is 3θ. Inserting eqs (1), (2) into eq. (6), we obtain

$$\tfrac{3}{2} \cos^2 \theta - \tfrac{1}{2} \cos^2 3\theta \leq 1\,.$$

With the choice $\theta = \pi/6$, we evaluate the left-hand side as $\frac{9}{8}$, showing indeed that quantum mechanics predicts a violation of the Bell inequality eq. (6). This prediction has been confirmed in the experiments of Aspect and of others.

4. Noncontextuality and Bell's inequality

A different derivation of the inequality starts from the supposition that separate joint probabilities exist for all of the four combinations of polarizer settings. Let us now write $P(a_+, b_+)$ in place of $P(a, b)$, to emphasize that this is the probability of *transmission* at A and B, with the polarizer settings a and b respectively. The corresponding probability for *absorption* at A and B is written $P(a_-, b_-)$, while $P(a_+, b_-)$ and $P(a_-, b_+)$ are the probabilities for transmission at one polarizer and absorption at the other. We set

$$P(a_+, b_+) = P(a_+, a'_+, b'_+, b_+) + P(a_+, a'_+, b'_-, b_+) \\ +P(a_+, a'_-, b'_+, b_+) + P(a_+, a'_-, b'_-, b_+) \tag{7}$$

which might be given the following Kolmogorovian, counterfactual interpretation. Consider the set of pairs of photons that are transmitted, one at A and one at B, when the settings are respectively a and b. Imagine that this set is divided into four disjoint sets, namely the subset that, if the settings had been a' and b' at A and B, both photons would have been transmitted, or such that transmission at A and absorption at B would have taken place, or absorption at A and transmission at B, or finally absorption at A and at B. Axiom V of Kolmogorov must be invoked to justify the addition of probabilities for these exclusive situations. It is supposed that each photon pair has, *at the same time*, the proclivity to be transmitted if a and b are the settings, *and* one or other of the four exclusive proclivities with respect to the counterfactual settings a' and b'.

Noncontextuality means that, for example, if the settings really are a' and b', instead of a and b, then the corresponding joint probability can now be divided into the following counterfactual subsets:

$$P(a'_+, b'_+) = P(a_+, a'_+, b'_+, b_+) + P(a_+, a'_+, b'_+, b_-) \\ +P(a_-, a'_+, b'_+, b_+) + P(a_-, a'_+, b'_+, b_-)\,.$$

Here the first term on the right is supposed to be the same as the first term on the right of eq. (7). That is, the counterfactual probability that a photon pair would have been transmitted if the settings had been a and b, given that they are a' and b', is the same as the corresponding probability if the settings had been a' and b', given that they are a and b (and similarly for all the other possible combinations). This assumption is natural from Einstein's realist viewpoint: the idea would be that a given pair of photons either does, or does not have the necessary properties to ensure transmission when the settings are *either* a and b or a' and b'. On the other hand, the assumption would have been anathema to Bohr, for whom the proclivities are joint properties of the photons and of the macroscopic measuring system. The choice of a and b for the settings specifies one macroscopic measuring system, and the choice of a' and b' specifies another. For him the counterfactual probabilities have no meaning, since if the photons are detected with one setting, they cannot be detected with another. The following derivation of the Bell inequality from the assumption of noncontextuality, together with the violation of the inequality in the Aspect experiment, supports Bohr's view at the expense of Einstein's Weltanschauung.

Let us streamline the notation before proceeding further. We write P_{++} in place of $P(a_+, b_+)$ and ρ_{+jk+} for the four probabilities on the right-hand side of eq. (7), where j and k can take on the values $\pm$. Consider

$$P_{i\ell} = \sum_{jk} \rho_{ijk\ell} \, . \tag{8}$$

Here i and ℓ go over $\pm$, and the case of eq. (7) corresponds to $i = +$ and $\ell = +$. We have here four probabilities, $P_{\pm\pm}$, and sixteen counterfactual conditional probabilities $\rho_{ijk\ell}$. In accordance with Kolmogorov's axiom III, all these probabilities are non-negative, the ρ as well as the P.

We may consider, instead of the above, three alternative cases. First, if the settings are a at A, but b' at B, then the probabilities are $P(a_\pm, b'_\pm)$, and we shall write them as $Q_{\pm\pm}$. Evidently

$$Q_{ik} = \sum_{j\ell} \rho_{ijk\ell} \, . \tag{9}$$

Similarly, with a' at A, and b at the B, the probabilities are

$$R_{j\ell} = \sum_{ik} \rho_{ijk\ell} \, , \tag{10}$$

where $R_{\pm\pm} = P(a'_\pm, b_\pm)$. Finally, with a' at A, and b' at the B, the probabilities are

$$S_{jk} = \sum_{i\ell} \rho_{ijk\ell} \, , \tag{11}$$

where $S_{\pm\pm} = P(a'_\pm, b'_\pm)$.

It is easy to check the following expression for the Bell coefficient:

$$\begin{aligned} B &\equiv P_{++} + Q_{++} + R_{++} - S_{++} \\ &= 2\rho_{++++} + 2\rho_{++-+} + 2\rho_{+-++} + \rho_{+--+} + \rho_{-+-+} - \rho_{-++-} . \end{aligned}$$

On the other hand, the probability that the photon has the proclivity to be transmitted with the setting a at A, irrespective of what happens at B, is

$$P(a_+) = P(a_+, b_+) + P(a_+, b_-) = \sum_{jk\ell} \rho_{+jk\ell} \, .$$

Similarly, the probability that the photon has the proclivity to be transmitted with the setting b at B, irrespective of what happens at A, is

$$P(b_+) = P(a_+, b_+) + P(a_-, b_+) = \sum_{ijk} \rho_{ijk+} \, .$$

A short calculation yields

$$\begin{aligned} P(a_+) + P(b_+) - B = {} & \rho_{+--+} + \rho_{+++-} + \rho_{++--} + \rho_{+-+-} + \rho_{+---} \\ & + \rho_{-+++} + \rho_{--++} + \rho_{---+} + \rho_{-++-} . \end{aligned}$$

This is non-negative, since none of the $\rho_{ijk\ell}$ are negative. In terms of the original notation of eq. (7), we have shown that

$$B = P(a_+, b_+) + P(a_+, b'_+) + P(a'_+, b_+) - P(a'_+, b'_+) \leq P(a_+) + P(b_+) \, ,$$

which is the Bell inequality, agreeing with eq. (6) in the notation of this section. It has been shown to be a consequence of the assumed existence of (noncontextual) joint probabilities that satisfy the Kolmogorov axioms.

5. Representation theorem

In this section, we start with the sixteen probabilities, P_{ij}, Q_{ij}, R_{ij}, S_{ij}, subject to the normalization conditions

$$\begin{aligned} P_{++} + P_{+-} + P_{-+} + P_{--} &= 1 = Q_{++} + Q_{+-} + Q_{-+} + Q_{--} \\ R_{++} + R_{+-} + R_{-+} + R_{--} &= 1 = S_{++} + S_{+-} + S_{-+} + S_{--} \,. \end{aligned} \tag{12}$$

The question is whether these quantities admit the representation eqs (8)–(11), with 16 positive weights $\rho_{ijk\ell}$ (noncontextual conditional probabilities). In the first place, the answer is certainly *no*, unless, in addition to the normalization conditions, the following constraints are satisfied:

$$\begin{aligned} P_{++} + P_{+-} &= Q_{++} + Q_{+-} \qquad & P_{++} + P_{-+} &= R_{++} + R_{-+} \\ S_{++} + S_{+-} &= R_{++} + R_{+-} \qquad & S_{++} + S_{-+} &= Q_{++} + Q_{-+} \,. \end{aligned} \tag{13}$$

This may be seen by writing out the left- and right-hand sides of eq. (13) in terms of the ρ : one finds identities. Note that no use is made at this stage of the positivity of the ρ.

So let us restate the question: given that the positive P, Q, R and S satisfy the constraints eq. (12) and eq. (13), is there always a representation of the form eqs (8)–(11) in which all the ρ are positive? We shall show that, if we drop the requirement that the ρ are positive, then there is indeed a solution, but it is not unique. Moreover, for some P, Q, R and S, we shall show that there are no solutions for which all the ρ are non-negative. That this must be so follows from the fact that the Bell inequality is violated for some P, Q, R and S, whereas if the ρ were positive in such cases, one could derive that inequality.

Let us first ask the restricted question: is it possible always to find $\rho_{ijk\ell}$ if we only specify the P, Q and R as given, positive quantities, obeying those of the constraints that do not involve the S? That this *is* possible we now show by construction. Consider

$$\rho_{ijk\ell} = \frac{P_{i\ell} Q_{ik} R_{j\ell}}{Q_i R_\ell} \,. \tag{14}$$

Here

$$Q_i = \sum_k Q_{ik} = \sum_\ell P_{i\ell}$$

and

$$R_\ell = \sum_j R_{j\ell} = \sum_i P_{i\ell} \,,$$

i.e.

$$\begin{aligned} Q_+ = P_{++} + P_{+-} &= Q_{++} + Q_{+-} \\ Q_- = P_{-+} + P_{--} &= Q_{-+} + Q_{--} \\ R_+ = P_{++} + P_{-+} &= R_{++} + R_{-+} \\ R_- = P_{+-} + P_{--} &= R_{+-} + R_{--} \,. \end{aligned}$$

Note that these expressions are consistent with the constraints eq. (12) and eq. (13). Clearly the realization eq. (14) yields non-negative $\rho_{ijk\ell}$, since all the probabilities are positive. We shall now show that eqs (8)–(10) are respected:

$$\sum_{jk} \rho_{ijk\ell} = P_{i\ell} \sum_k \frac{Q_{ik}}{Q_i R_\ell} \sum_j R_{j\ell} = P_{i\ell} \sum_k \frac{Q_{ik}}{Q_i} = P_{i\ell}$$

$$\sum_{j\ell} \rho_{ijk\ell} = Q_{ik} \sum_\ell \frac{P_{i\ell}}{Q_i R_\ell} \sum_j R_{j\ell} = Q_{ik} \sum_\ell \frac{P_{i\ell}}{Q_i} = Q_{ik}$$

$$\sum_{ik} \rho_{ijk\ell} = R_{j\ell} \sum_i \frac{P_{i\ell}}{Q_i R_\ell} \sum_k Q_{ik} = R_{j\ell} \sum_i \frac{P_{i\ell}}{R_\ell} = R_{j\ell}.$$

This concludes the demonstration.

The above construction shows that, for any acceptable P, Q and R, there is a representation of the required form. However, although the four probabilities S_{jk} could be calculated from eq. (11), using the ρ that have been constructed, there is no guarantee that they would agree with the S that are given (or measured). To complete the existence proof, and to investigate the question of uniqueness, we turn to matrix theory.

6. Reduction to matrix form

In this section we rewrite the above equations in matrix form. The purpose is to use the general methods of matrix algebra to study the most general solution of our problem. We define

$$\begin{bmatrix} P_{++} \\ P_{+-} \\ P_{-+} \\ P_{--} \\ Q_{++} \\ Q_{+-} \\ Q_{-+} \\ Q_{--} \\ R_{++} \\ R_{+-} \\ R_{-+} \\ R_{--} \\ S_{++} \\ S_{+-} \\ S_{-+} \\ S_{--} \end{bmatrix} = \begin{bmatrix} 1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&1&1&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1 \\ 1&0&1&0&1&0&1&0&0&0&0&0&0&0&0&0 \\ 0&1&0&1&0&1&0&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&1&0&1&0&1&0 \\ 0&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1 \\ 1&1&0&0&0&0&0&0&1&1&0&0&0&0&0&0 \\ 0&0&0&0&1&1&0&0&0&0&0&0&1&1&0&0 \\ 0&0&1&1&0&0&0&0&0&0&1&1&0&0&0&0 \\ 0&0&0&0&0&0&1&1&0&0&0&0&0&0&1&1 \\ 1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0 \\ 0&1&0&0&0&1&0&0&0&1&0&0&0&1&0&0 \\ 0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0 \\ 0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1 \end{bmatrix} \begin{bmatrix} \rho_{++++} \\ \rho_{++-+} \\ \rho_{+-++} \\ \rho_{+--+} \\ \rho_{+++-} \\ \rho_{++--} \\ \rho_{+-+-} \\ \rho_{+---} \\ \rho_{-+++} \\ \rho_{-+-+} \\ \rho_{--++} \\ \rho_{---+} \\ \rho_{-++-} \\ \rho_{-+--} \\ \rho_{--+-} \\ \rho_{----} \end{bmatrix}$$

or for short

$$F = Mx\,, \tag{15}$$

where F is the vector of the P, Q, R and S on the left-hand side, where M is the square matrix of elements 1 and 0, and where x is the vector of the ρ.

The matrix M is singular, its rank being in fact only 4. There are 7 independent eigenvectors belonging to zero eigenvalue:

$$Mu_j = 0\,,$$

for $j = 1, 2, \ldots, 7$. These vectors u_j span the seven-dimensional null space of the matrix M. There are also four nonzero eigenvalues:

$$Mw_j = \lambda_j w_j\,,$$

for $j = 1, 2, 3, 4$, where $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 2$, $\lambda_4 = 4$.

Eleven independent (but not orthonormalized) eigenvectors corresponding to these eigenvalues are as follows:

$$\begin{array}{ccccccc} u_1 & u_2 & u_3 & u_4 & u_5 & u_6 & u_7 \\ \begin{bmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\\-1\\0\\0\\-1\\1 \end{bmatrix} & \begin{bmatrix} 0\\0\\0\\0\\1\\0\\-1\\0\\0\\0\\0\\0\\-1\\0\\1\\0 \end{bmatrix} & \begin{bmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\1\\-1\\0\\0\\-1\\1\\0\\0 \end{bmatrix} & \begin{bmatrix} 0\\1\\0\\-1\\0\\0\\0\\0\\0\\-1\\0\\1\\0\\0\\0\\0 \end{bmatrix} & \begin{bmatrix} 1\\0\\-1\\0\\0\\0\\0\\0\\-1\\0\\1\\0\\0\\0\\0\\0 \end{bmatrix} & \begin{bmatrix} 0\\0\\1\\-1\\0\\0\\-1\\1\\0\\0\\0\\0\\0\\0\\0\\0 \end{bmatrix} & \begin{bmatrix} 1\\-1\\0\\0\\-1\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{bmatrix} \end{array}$$

$$\begin{array}{cccc} w_1 & w_2 & w_3 & w_4 \\ \begin{bmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\\-1\\-1\\1 \end{bmatrix} & \begin{bmatrix} 0\\0\\0\\0\\-1\\1\\-1\\1\\-1\\-1\\1\\1\\-2\\0\\0\\2 \end{bmatrix} & \begin{bmatrix} 0\\0\\0\\0\\1\\-1\\1\\-1\\-1\\-1\\1\\1\\0\\-2\\2\\0 \end{bmatrix} & \begin{bmatrix} 1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1 \end{bmatrix} \end{array}\,.$$

Since M is singular, eq. (15) does not have solutions for arbitrary F; the necessary and sufficient condition that solutions exist is the orthogonality of F to the null space of M^T, i.e. F must satisfy

$$v_j F = 0\,, \tag{16}$$

for $j = 1, 2, \ldots, 7$, where the v_j span the seven-dimensional adjoint null space:

$$v_j M = 0\,.$$

This null space is spanned by the eigenvectors

$$\begin{aligned}
v_1 &= [\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,-1,\ \ 0,-1,\ \ 0,\ \ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 1\ ],\\
v_2 &= [\ \ 0,\ \ 0,\ \ 0,\ \ 0,-1,\ \ 0,-1,\ \ 0,\ \ 0,\ \ 0,\ 0,\ 0,\ 1,\ 0,\ 1,\ 0\ ],\\
v_3 &= [\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,-1,-1,\ 0,\ 0,\ 1,\ 1,\ 0,\ 0\ ],\\
v_4 &= [\ \ 0,-1,\ \ 0,-1,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 1,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0\ ],\\
v_5 &= [-1,\ \ 0,-1,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 1,\ \ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0\ ],\\
v_6 &= [\ \ 0,\ \ 0,-1,-1,\ \ 0,\ \ 0,\ \ 1,\ \ 1,\ \ 0,\ \ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0\ ],\\
v_7 &= [-1,-1,\ \ 0,\ \ 0,\ \ 1,\ \ 1,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0\ ].
\end{aligned}$$

On writing out eq. (16) and performing a little algebra, we find that these constraints are equivalent to the requirements eq. (12) and eq. (13) (except that the normalization is arbitrary). This is a satisfactory state of affairs: the necessary constraints are also sufficient conditions for the solubility of the matrix equations.

Suppose now that we have an F that respects the constraints eq. (16), and suppose that x is a particular solution of eq. (15). The most general solution, corresponding to this particular F, has the form

$$x + \sum_{j=1}^{7} \alpha_j u_j\,,$$

for any real coefficients α_j.

7. Construction of probabilities

After this detour into matrix theory, we return to our representation theorem. We have seen that it is always possible to construct a set of $\rho_{ijk\ell}$ that fits any specified, acceptable set of probabilities, $P_{i\ell}$, Q_{ik} and $R_{j\ell}$, but that the corresponding values of S_{jk} are not guaranteed to be as specified. Suppose that we write the ρ that we have constructed as a vector, x, in the manner of eq. (15). What must we add to x to rectify the S values? Evidently we must add something with care, for the P, Q and R are already correct and so must not be disturbed. Since any F that satisfies eq. (15) necessarily satisfies also the constraints eq. (12) and eq. (13), if we change the S while leaving the P, Q and R unchanged, the changes in the S must satisfy

$$\begin{aligned}
\Delta(S_{++} + S_{+-}) &= 0,\\
\Delta(S_{++} + S_{-+}) &= 0,\\
\Delta(S_{++} + S_{+-} + S_{-+} + S_{--}) &= 0\,,
\end{aligned} \tag{17}$$

which implies

$$\Delta S_{++} = -\Delta S_{+-} = -\Delta S_{-+} = \Delta S_{--} \, .$$

The most general change in F that is allowed is therefore a constant multiple of w_1, the eigenvector of M belonging to the eigenvalue unity, for we recall that

$$w_1^T = [\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ -1,\ -1,\ 1\ \ ] \, .$$

Note that w_1 is a 1×16 vertical array: we have displayed its transpose to save space on the page, and we will resort to this stratagem again below. We write the modified vector as

$$y = x + \beta w_1 \, , \tag{18}$$

and note that

$$Fy = Fx + \beta w_1 \, .$$

Hence the addition of a multiple of w_1 succeeds indeed in changing only the S_{jk}, i.e. only F_{13}, F_{14}, F_{15} and F_{16}, and moreover in the only way that is consistent with the restrictions that the S_{jk} must satisfy. It is then enough to choose β such that one of the S_{jk} is as specified. For example, choose

$$\beta = F_{13} - x_1 - x_5 - x_9 - x_{13} \, , \tag{19}$$

thus guaranteeing that $S_{++} = F_{13}$ is correct, and therefore also the remaining S. Although we have succeeded in fitting the specified F, there is no guarantee that all the components of y are non-negative, since w_1 has positive and negative components. Moreover, we know that the solution is not unique, for M has a seven dimensional null-space.

The most general solution of the representation problem, ignoring the requirement of positivity, is

$$z = x + \beta w_1 \ + \sum_{j=1}^{7} \alpha_j u_j \, , \tag{20}$$

where x corresponds to the $\rho_{ijk\ell}$ as constructed in eq. (14), where β is given by eq. (19), and where the α_j are seven arbitrary real numbers.

8. Negative probabilities

As in §3, consider the case in which the angle between the directions of a and b, between a and b', and between a' and b is $\pi/6$, while that between a' and b' is thrice that, i.e. $\pi/2$. Then

$$\begin{aligned} P_{++} = P_{--} = Q_{++} = Q_{--} = R_{++} = R_{--} &= \tfrac{3}{8}, \\ P_{+-} = P_{+-} = Q_{+-} = Q_{+-} = R_{+-} = R_{-+} &= \tfrac{1}{8}, \\ S_{++} = S_{--} &= 0, \\ S_{+-} = S_{-+} &= \tfrac{1}{2}. \end{aligned} \tag{21}$$

In terms of the notation of the previous section, we can write these results in matrix form as follows:

$$F^T = \tfrac{1}{8}\,[\ 3,\ 1,\ 1,\ 3,\ 3,\ 1,\ 1,\ 3,\ 3,\ 1,\ 1,\ 3,\ 0,\ 4,\ 4,\ 0\ \ ]. \qquad (22)$$

We next use the formula eq. (14) to calculate a $\rho_{ijk\ell}$ that reproduces $P_{i\ell}$, Q_{ik}, $R_{j\ell}$, but not S_{jk}:

$$\begin{aligned} \rho_{++++} = \rho_{----} &= \tfrac{27}{128}, \\ \rho_{++-+} = \rho_{+-++} = \rho_{+-+-} = \rho_{-+-+} = \rho_{-+--} = \rho_{--+-} &= \tfrac{9}{128}, \\ \rho_{+--+} = \rho_{+---} = \rho_{---+} = \rho_{-++-} &= \tfrac{3}{128}, \\ \rho_{++--} = \rho_{--++} &= \tfrac{1}{128}, \end{aligned} \qquad (23)$$

or, in terms of the notation of eq. (15),

$$x^T = \tfrac{1}{128}\,[\ 27,\ 9,\ 9,\ 3,\ 3,\ 1,\ 9,\ 3,\ 3,\ 9,\ 1,\ 3,\ 3,\ 9,\ 9,\ 27\ \ ]. \qquad (24)$$

On inserting this x into eq. (15), we obtain

$$F^T = \tfrac{1}{8}\,[\ 3,\ 1,\ 1,\ 3,\ 3,\ 1,\ 1,\ 3,\ 3,\ 1,\ 1,\ 3,\ \tfrac{9}{4},\ \tfrac{7}{4},\ \tfrac{7}{4},\ \tfrac{9}{4}\ \ ]. \qquad (25)$$

Comparing this with eq. (22), we see that the components 1–12 are correct, for these correspond to the P, the Q and the R, but the remaining four components, which correspond to the S, are wrong.

Now we see from eq. (25) that $F_{13} = \frac{9}{32}$, whereas it should be 0. Accordingly, we must subtract $\frac{9}{32}w_1$ from x, as given in eq. (8.), for this will replace the F of eq. (25) by $F - \frac{9}{32}w_1$, which is

$$\tfrac{1}{8}\,[\ 3,\ 1,\ 1,\ 3,\ 3,\ 1,\ 1,\ 3,\ 3,\ 1,\ 1,\ 3,\ 0,\ 4,\ 4,\ 0\ \ ],$$

agreeing indeed with eq. (22). However, the price is that the new expression for x is

$$x^T = \tfrac{1}{128}\,[\ 27,\ 9,\ 9,\ 3,\ 3,\ 1,\ 9,\ 3,\ 3,\ 9,\ 1,\ 3,\ -33,\ 45,\ 45,\ -9\ \ ].$$

As can be seen, x_{13} and x_{16}, i.e. ρ_{-++-} and ρ_{----}, are negative, which is inconsistent with their interpretation as probabilities.

Is there any way to remove the negativity of these components of x without spoiling the fit to F? There is not, for the most general solution is to replace x by z, see eq. (8.) and eq. (20), with $\beta = -\frac{9}{32}$ and arbitrary real α_j. In the Appendix, we show that there is no choice of the α's such that all the components of z are non-negative. This is a direct demonstration of what we already know indirectly, for on the one hand we have shown that the Bell inequality can be proved if none of the ρ are negative, and on the other hand we know that the inequality is in fact violated for the choice of angles in question.

Appendix

Here we show that the most general form, z, as constructed in §8, is such that at least one of its components is negative. Thus there is no solution with non-negative components.

To effect the proof, we assume on the contrary that there is a choice of the coefficients, α_j , $j = 1, 2, ..., 7$, such that $z_n \geq 0$, $n = 1, 2, ..., 16$, and we will obtain a contradiction, thus concluding the proof by *reductio ad absurdum.*

For convenience, we replace the α by new coefficients, γ, defined by

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \\ \gamma_6 \\ \gamma_7 \end{bmatrix} = \begin{bmatrix} -9 \\ 45 \\ 45 \\ 9 \\ 1 \\ 3 \\ 1 \end{bmatrix} + 128 \begin{bmatrix} \alpha_1 \\ \alpha_2 - \alpha_1 \\ \alpha_3 \\ \alpha_4 - \alpha_7 \\ \alpha_1 + \alpha_5 \\ \alpha_6 \\ \alpha_7 \end{bmatrix} . \tag{26}$$

In terms of the γ_j, we have

$$z = \frac{1}{128} \begin{bmatrix} 16 - \gamma_1 + \gamma_5 + \gamma_7 \\ \gamma_4 \\ 16 + \gamma_1 - \gamma_5 + \gamma_6 \\ 16 - \gamma_4 - \gamma_6 - \gamma_7 \\ -32 + \gamma_1 + \gamma_2 - \gamma_7 \\ \gamma_7 \\ 48 - \gamma_1 - \gamma_2 - \gamma_6 \\ \gamma_6 \\ -32 + \gamma_1 + \gamma_3 - \gamma_5 \\ 64 - \gamma_3 - \gamma_4 - \gamma_7 \\ \gamma_5 \\ -16 - \gamma_1 + \gamma_4 + \gamma_7 \\ 48 - \gamma_1 - \gamma_2 - \gamma_3 \\ \gamma_3 \\ \gamma_2 \\ \gamma_1 \end{bmatrix} \tag{27}$$

and it is clear that all the γ_j must be non-negative: only in this way can the second, sixth, eighth, eleventh, fourteenth, fifteenth and sixteenth components of z be non-negative.

If all of the components of z are non-negative, the sum of any two components is also non-negative. In particular,

$$128(z_4 + z_{12}) = -\gamma_1 - \gamma_6 \geq 0 \, ,$$

but since the γ_j are non-negative, we must have

$$\gamma_1 = 0 = \gamma_6 \, .$$

It now follows that

$$128(z_5 + z_9 + z_{13}) = -16 - \gamma_5 - \gamma_7 \, ,$$

but this is impossible, since the left-hand side is non-negative, but the right-hand side cannot be greater than -16. This incompatibility constitutes the contradiction that we sought.

Acknowledgement

I would like to thank S M Roy for discussions, and for the hospitality extended to me over the years at TIFR, Mumbai.

References

[1] Bas van Fraassen, *Quantum mechanics, an empiricist view* (Clarendon Press, 1991) p. 112
[2] B Kümmerer and H Maasen, Elements of quantum probability, in *Quantum Probability Communications X* edited by R L Hudson and J M Lindsay (World Scientific, Singapore, 1996)
[3] R Gill, Critique of elements of quantum probability, in *Quantum Probability Communications XI* edited by R L Hudson and J M Lindsay (World Scientific, Singapore, 1996)
[4] For that matter, the Copernican heliocentric theory is also not *required* by the facts of planetary motions. A geocentric, geostationary coordinate system may be used, and planetary and solar motions with respect to such axes may be expanded in Fourier-Ptolemy series of epicycles
[5] A Fine, *Phys. Rev. Lett.* **48**, 291 (1982)
[6] A Aspect, J Dalibard and G Roger, *Phys. Rev. Lett.* **49**, 1804 (1982)
[7] D Atkinson, *Dialectica* **52**, 103 (1998)
[8] David Atkinson and Jeanne Peijnenburg, *Synthese* **118**, 307 (1999)
[9] A N Kolmogorov, *Foundations of the theory of probability*, translation, edited by N Morrison (Chelsea Publishing Company, New York, 1956)