



**HAL**  
open science

# A logical loophole in the derivation of the Bell inequalities

Gerrit Coddens

► **To cite this version:**

| Gerrit Coddens. A logical loophole in the derivation of the Bell inequalities. 2018. cea-01737341v6

**HAL Id: cea-01737341**

**<https://cea.hal.science/cea-01737341v6>**

Preprint submitted on 1 May 2018 (v6), last revised 19 Sep 2018 (v7)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A logical loophole in the derivation of the Bell inequalities

Gerrit Coddens

Laboratoire des Solides Irradiés,  
Université de Paris-Saclay, CEA-DRF-IRAMIS, CNRS UMR 7642, Ecole Polytechnique,  
28, Route de Saclay, F-91128-Palaiseau CEDEX, France

1st May 2018

**Abstract.** The Bell inequalities are based on a tacit assumption of a *common* probability distribution that precludes their application to the experiments of Aspect *et al.* The basic ideas of this argument have already been given in references [1,2], but the present presentation recollects them in a clearer and more concise way.<sup>a</sup> We show that the intuitive idea that it would be easy enough to define such a common probability distribution (and this way make up for the hiatus in the traditional proof) is wrong. We discuss two such intuitive repair scenarios, one based on Cartesian products of sets and one based on unions of sets. Both lead to an impasse for reasons that initially might have escaped our attention. When we change strategies to avoid an error in the first scenario, another error pops up in an entirely different place within the second scenario. Perhaps it is possible to invent yet other scenarios, such that the polemics could go on. But it cannot be anybody's task to imagine all possible attempts to repair the logical loophole and then rebut them one by one. That would be a reversal of the charge of proof. In the situation as it stands, the proof of the Bell inequalities is wrong due to the absence of an existence proof for the tacitly assumed common probability distribution. We discuss also a number of mathematical and physical hints which strongly suggest that the error is absolutely lethal and that no repair scenario will ever be able to provide us with the missing existence proof. This logical loophole is much more fundamental than the experimental loopholes which have currently received much interest due to the ingenious efforts of experimentalists to close them.

**PACS.** 03.65.-w Quantum Mechanics

## 1 The Bell inequalities and their application to the experiments of Aspect *et al.*

The subject matter of the Bell inequalities and the experiments of Aspect *et al.* hardly needs any introduction [4]. However, the argument has often been blurred by drawing in unnecessary issues, leading to some confusion. We give here an elementary derivation that removes all unnecessary considerations. This will show how elementary the argument is and how very hard it is to question the validity of the inequalities. This may perhaps also help in convincing the reader that I am not “one of those cranks” again such that it could be worth to continue reading.

We consider 4 variables  $a_1 \in S$ ,  $a_2 \in S$ ,  $b_1 \in S$ ,  $b_2 \in S$ , where  $S = \{0, 1\}$ . The idea is that 0 corresponds to absorption in a polarizer, and 1 to transmission.  $a_j$  will correspond to polarizer settings in one arm of the set-up,  $b_k$  to polarizer settings in the other arm. There are thus 16 possible combinations for the values of  $(a_1, a_2, b_1, b_2)$ . By making a table of these 16 combinations it is easy to verify that we always have:

$$\forall (a_1, a_2, b_1, b_2) \in S^4 : \quad Q = a_1 b_1 - a_1 b_2 - a_2 b_1 - a_2 b_2 + a_2 + b_2 \in S. \quad (1)$$

We consider now functions  $a_j \in F(V, S)$  and  $b_k \in F(V, S)$ . The notation means that the domain of the functions is  $V$ , while the functions take their values in  $S$ . Here  $V$  is a set of relevant variables for the experiment. We can call the set  $V$  the set of hidden variables, even if some of them may not really be hidden. One can imagine that  $V$  could be a subset of a vector space  $\mathbb{R}^n$  or of a manifold, e.g. a non-abelian Lie group like  $\text{SO}(3)$ . We have then:

$$\forall \lambda \in V : \quad 0 \leq Q(\lambda) = a_1(\lambda)b_1(\lambda) - a_1(\lambda)b_2(\lambda) - a_2(\lambda)b_1(\lambda) - a_2(\lambda)b_2(\lambda) + a_2(\lambda) + b_2(\lambda) \leq 1. \quad (2)$$

---

<sup>a</sup> The paper supersedes reference [3] which contains an error. It also supersedes its own previous versions because certain parts were not sufficiently clearly explained and some of the considerations were not really needed. We are sorry for the inconvenience.

We can now consider a probability density  $p$  over  $V$ , i.e.  $p(\lambda) d\lambda$ . The function  $p$  belongs then to the set of functions  $F(V, [0, \infty[)$  with domain  $V$  and values in  $[0, \infty[$ . We further require that  $\int_V p(\lambda) d\lambda = 1$ . We can now integrate Eq. 2 with  $p$  over  $V$ . Introducing the notations:

$$p(\alpha_j \wedge \beta_k) = \int_V a_j(\lambda) b_k(\lambda) p(\lambda) d\lambda, \quad p(\alpha_j) = \int_V a_j(\lambda) p(\lambda) d\lambda, \quad p(\beta_k) = \int_V b_k(\lambda) p(\lambda) d\lambda, \quad (3)$$

we obtain then:

$$0 \leq p(\alpha_1 \wedge \beta_1) - p(\alpha_1 \wedge \beta_2) - p(\alpha_2 \wedge \beta_1) - p(\alpha_2 \wedge \beta_2) + p(\alpha_2) + p(\beta_2) \leq 1. \quad (4)$$

This is the CHSH Bell inequality used in the experiments of Aspect et al. It is a purely mathematical identity and does not depend on any physical considerations. It is also free of any considerations about statistical correlations and statistical independence. The probabilities are identified with the mathematical expressions for the outcomes of the photon polarization experiments reported by Aspect *et al.*:

$$p(\alpha_j \wedge \beta_k) = \frac{1}{2} \cos^2(\alpha_j - \beta_k), \quad p(\alpha_j) = \frac{1}{2}, \quad p(\beta_k) = \frac{1}{2}, \quad (5)$$

where  $\alpha_j$  and  $\beta_k$  are the angles of the polarizer settings in the two arms of the experiment. According to quantum theory the mathematical expressions are the limits of the measured probabilities when the number of registered events tends to infinity, i.e. when the statistics become perfect. For a function  $f \in F(\mathbb{N}, \mathbb{R})$ , the limit  $n \rightarrow \infty$  is defined by:

$$\lim_{n \rightarrow \infty} f(n) = F \Leftrightarrow (\forall \varepsilon > 0)(\exists N \in \mathbb{N})(n > N \Rightarrow |f(n) - F| < \varepsilon). \quad (6)$$

Here  $f(n)$  would be the measured probabilities after  $n$  detection events,  $F$  the theoretical expression  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$ , and  $\varepsilon$  the statistical accuracy of the experiment required. An experimentalist has to worry about the statistical and instrumental precisions. For practical reasons the experimentalist can only reach a reasonable accuracy  $\varepsilon$ . But this should be well enough to establish beyond any reasonable doubt if the Bell inequality is satisfied or otherwise. We will adopt a mathematician's viewpoint and assume that the expressions  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$  are exact, trusting that at least in principle the experimentalist could prove this to any accuracy  $\varepsilon$ , by improving the experimental protocol. We introduce thus the assumption (or act of faith) that the algebra of quantum mechanics is exact. This frees us from all imaginable polemics about experimental bias. For certain values of  $(\alpha_1, \alpha_2, \beta_1, \beta_2) \in [0, 2\pi]^4$ , the expressions in Eq. 5 do not satisfy the inequality in Eq. 4. This violation of the Bell inequality shows that the mathematical expressions in Eq. 5 are not compatible with Eq. 4. This seems to confirm Bohr's thesis that the polarizations cannot exist prior to a measurement and must be created by the measurement. But it is then extremely puzzling that we can obtain a definite correlation  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$  because the polarizers can be separated by arbitrarily large distances. It looks like the spooky action at a distance Einstein talked about and which has been called entanglement in the aftermath of the experiments. In the experiments of Aspect this issue is tested by ensuring Einstein separation of the detection events in both arms. The solution of this conundrum is in our opinion summarized in the last paragraph of Subsection 4.3 and the last sentence of Subsection 4.4.

## 2 The tacit assumption

The derivation of the inequality looks unassailable. It is indeed ought to be too simple to possibly hide a logical loophole. But it does! What is not acknowledged is that the identification in Eq. 5 introduces a tacit assumption, which is admittedly hard to discern, namely that all four quantities  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$  can be obtained from an integration over a same set  $V$  with a *same common distribution function*  $p$ . However, it can *a priori* not be excluded that in reality we can only obtain the quantities  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$  from four different distributions  $p_{jk}$  according to:

$$\frac{1}{2} \cos^2(\alpha_j - \beta_k) = \int_{V_{jk}} a_j(\lambda) b_k(\lambda) p_{jk}(\lambda) d\lambda \quad (7)$$

rather than:

$$\exists V, \exists! p \in F(V, [0, \infty[) \parallel \forall (j, k) \in \{1, 2\}^2 : \frac{1}{2} \cos^2(\alpha_j - \beta_k) = \int_V a_j(\lambda) b_k(\lambda) p(\lambda) d\lambda. \quad (8)$$

In other words, it is tacitly assumed that the quantities  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$  can all be obtained from *one single common distribution function*  $p$  rather than from four different distributions  $p_{jk}$ . In view of the importance of the subject matter, one may desire an existence proof of such a unique function  $p$ . For a mathematician it will be immediately obvious that the necessity of such an existence proof is compelling. Unfortunately, a physicist may remain unfazed by

a request for such a proof. It just would not set his alarm bells ringing. He may take the existence for self-evident and consider the request as unproductive and faultfinding mathematical nitpicking.

It is obvious that if we turn a polarizer, we turn a distribution of molecules. If one believes in hidden variables, then this distribution of molecules must be part of the hidden variables, because  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$  changes when we turn one of the polarizers. Hence by turning a polarizer we change the hidden variables, such that *there are indeed four different sets*  $V_{jk}$ . One may have the intuition that it is trivial to make up for this. All one would have to do is to enlarge the sets  $V_{jk}$  to a set  $V$ , where  $\forall (j, k) \in \{1, 2\}^2 : V_{jk} \subset V$ . But the problem is not only that we have to construct  $V$ . There is also a normalization problem as *the measured quantities*  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$  *are beyond any doubt normalized individually according to*  $\int_{V_{jk}} p_{jk}(\lambda) d\lambda = 1$  *rather than globally according to*  $\int_V p(\lambda) d\lambda = 1$  (if  $V$  exists). In fact, the three probability distributions  $p_{j'k'}$  for  $(j', k') \neq (j, k)$  do not intervene in the measurement of  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$ , which stands on its own and is independent of our intentions to consider other experiments with the aim of measuring four quantities in total. This normalization problem may well be the reason for the violation of the Bell inequality. We will work this out in Section 4 where we will also discuss an approach based on cartesian products of sets rather than on unions of sets..

### 3 Discussion: Why the objection is not at all futile or contrived

It is not my task to prove that assumption Eq. 8 is wrong. That would be a reversal of the charge of proof. All charge of proof is with the authors who proposed the Bell inequalities. I could stop here and wish them ironically good luck. However, I think it will be a more reasonable and respectful attitude to try to provide the physicists in the audience with some arguments why my objection is not as farfetched as they may think. These arguments may look like a blend of physics and mathematics, but we will show that on close inspection they reveal to be all purely mathematical.

- Our first argument is that the contents of our objection are vindicated by quantum mechanics itself. If one wanted to claim they are wrong, one would thus have to claim that quantum mechanics is wrong. This argument will lead us straight into the heart of the Bohr-Einstein debate. The starting point of this debate was that when two operators do not commute, they will not have common eigenvectors. This is a purely mathematical truth. An example of this are the angular momentum operators  $\hat{L}_x$  and  $\hat{L}_y$ .<sup>1</sup> According to Bohr the quantities  $L_x$  and  $L_y$  do then *not exist simultaneously*. Einstein wanted to prove that this cannot be true and proposed the EPR experiment for two correlated particles, whereby one would measure  $L_x$  for one of the particles at  $\mathbf{r}_1$  in one arm of the set-up, and  $L_y$  for the other particle at  $\mathbf{r}_2 = -\mathbf{r}_1$  in the other arm of the set-up. This would then demonstrate that Bohr was wrong. We know now that such a simultaneous measurement is indeed possible, because  $\hat{L}_{x_1} = y_1 \frac{\partial}{\partial z_1} - z_1 \frac{\partial}{\partial y_1}$  and  $\hat{L}_{y_2} = z_2 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial z_2}$  do commute when  $\mathbf{r}_1 \neq \mathbf{r}_2$ . It is only when  $\mathbf{r}_1 = \mathbf{r}_2$  that the operators do not commute, with all the consequences non-commutativity may entail. In a sense, Einstein addressed thus the wrong issue, because he no longer addressed operators that were not commuting, but he had been sidetracked by an overinterpretation of the mathematics introduced by Bohr.

We may in fact note that we discover here that the mathematical consequence of the fact that two operators  $\hat{O}_1$  and  $\hat{O}_2$  do not commute is in general not that  $O_1$  and  $O_y$  would not exist simultaneously as Bohr claimed, but that  $\hat{O}_1$  and  $\hat{O}_2$  do not have a common probability distribution function. In fact, this will be true if we can use the eigenfunctions  $\psi_j$  of  $\hat{O}_j \psi_j = O_j \psi_j$  to define the probability distributions  $p_j = |\psi_j|^2$  for a set-up used to measure  $O_j$ , based on the Born rule. It is not because  $O_1$  and  $O_2$  cannot be measured simultaneously in a same location that they would not exist simultaneously in the same location. The difference with what Bohr said may look perhaps subtle but it is very important, because it implies that Bohr has overinterpreted the mathematics. In the case of the operators  $\hat{L}_j$ , the whole issue is actually a matter which is not at all subject to interpretation, because it is all just pure mathematics, the mathematics of the rotation group. These mathematics provide the full explanation about what is going on with the operators  $\hat{L}_x$  and  $\hat{L}_y$  and their eigenfunctions. Bohr presumably did not know this group theory. Using his physical intuition he cooked up a parallel theory by guessing an *ad hoc* explanation. But physical intuition and mathematical intuition are different things. His guesses were at variance with the correct mathematics. They were therefore plain wrong, but nobody knew at that time that group theory was at stake and nobody noticed. Nobody should be blamed for this. Bohr then went on pushing his ideas forcefully. Our rebuttal of the Bell inequality pinpoints his overinterpretation very accurately. Bohr had a clash with Heisenberg over a very similar overinterpretation in a discussion about the uncertainty relations, which are also tied up with non-commuting operators.

Einstein reasoned on the overinterpretation provided to him by Bohr and was thus right with his intuition that this overinterpretation had to be wrong. We must further point out that the correct interpretation of the consequences of

<sup>1</sup> Mind that we do not claim that the operators  $L_j$  are playing a rôle in the experiments of Aspect *et al.*. We just use them as an example.

the fact that  $\hat{L}_x$  and  $\hat{L}_y$  do not commute can only be classical. The operators  $\hat{L}_x$  and  $\hat{L}_y$  exist in the group theory of the rotation group, which is mere Euclidean geometry. Their definitions and meaning are thus engraved in stone prior to any application in physics and are therefore not liable to any kind of warped alternative “physical” interpretation. Up to a proportionality constant  $-\hbar$ , the operator  $\hat{L}_z$  corresponds to  $\frac{\partial}{\partial\varphi}$  in spherical coordinates  $(r, \theta, \varphi)$  in  $\mathbb{R}^3$  or polar coordinates  $(r, \varphi)$  in the  $Oxy$  plane. The operators  $\hat{L}_j$  serve thus to calculate Lie derivatives in three mutually orthogonal directions. They are used to constitute a basis for the tangent space to the group at the identity element as they intervene in the calculation of the infinitesimal generators and of the elements of the Lie algebra. To obtain the Lie derivatives one must use one-parameter sets. These are just completely different for  $\hat{L}_x$  and  $\hat{L}_y$ . The fact that  $\hat{L}_x$  and  $\hat{L}_y$  do not have common eigenfunctions only means that a rotation around the  $x$ -axis cannot simultaneously be a rotation around the  $y$ -axis [1]. In fact,  $\hat{L}_x$  is not the operator for the  $x$ -component of the angular momentum  $\mathbf{L}$ , but for  $\mathbf{L}$  when it is aligned with the  $x$ -axis, such that then  $L_x = L$ ,  $L_y = 0$ ,  $L_z = 0$  and all components exist simultaneously [1] as they should in Euclidean geometry. This is illustrated by the fact that if we call  $R$  the rotation around the  $z$ -axis that rotates  $\mathbf{e}_x$  to  $\mathbf{e}_y$ , then  $\hat{L}_y = R\hat{L}_xR^{-1}$ . Hence  $\hat{L}_y$  is just the same Lie derivative as  $\hat{L}_x$  but in another direction, i.e. for another one-parameter set.<sup>2</sup> Furthermore, one can associate an uncertainty relation with the fact that  $\hat{L}_x$  and  $\hat{L}_y$  do not commute. But as said,  $\hat{L}_x$  and  $\hat{L}_y$  are just part of the representation theory of the rotation group and there is absolutely no uncertainty in Euclidean geometry. For sure, all this flies in the face of many things we have learned, but that is just because we have been duped with misinterpretations of the mathematics. In summary, the whole algebra of angular momentum belongs just to the theory of the representations of the rotation group in Euclidean geometry and as such to classical mechanics. Hence, even though we are as physicists introduced to the conceptual world of non-commuting operators and Lie groups by quantum mechanics, this world is not “quantum-mechanical” in the sense of non-classical. And to teach us these mathematics, our best guide may perhaps not be a quantum guru.

- Our second argument is another way of showing that the contents of our objection are vindicated by quantum mechanics itself. It is based on one of the ways one calculates probabilities in quantum mechanics. In fact, what one does is set up a Schrödinger equation, solve it to find the wave function  $\psi$  and then stipulate  $p = |\psi|^2$ , according to the Born rule. It must be mathematically obvious that one can in general not assume that the solution  $\psi_1$  of a first Schrödinger equation will be equal to the solution  $\psi_2$  of a second different Schrödinger equation. One can therefore in general not carry over probabilities that are valid for one set-up to another set-up. But this is exactly what the ansatz of a common probability distribution for the four different correlation experiments does. Of course this argument is completely equivalent to the first one. We just discover it here in a different guise.
- The first two arguments rely on the Born rule. This may come over as a cheat within a context where we are pitting quantum mechanics against classical thinking. We will explain below in Section 5 that the quadratic structure of the Born rule is not something specifically quantum mechanical but a very general result of group theory and therefore completely classical. The complete proof of the Born rule requires also physics and it can be proved in several cases that these are just classical physics. But at this stage the argument becomes too involved for what we need here, such that we abandoned the idea of developing it fully. The occurrence of various cases shows actually that there is not one Born rule, but several ones. To shortcut this problem we will therefore develop also classical arguments.
- A third argument which is completely classical beyond contention because it does not depend on the Born rule is that the definition of a probability depends on a full context and a protocol as one discovers by the paradox of Bertrand. One can connect this to the previous arguments. When we solve a Schrödinger equation, we take into account the necessity of outlining the full context by coding it into the boundary conditions, perhaps even unwittingly. It is well known that solutions of a Dirichlet problem can heavily depend on the boundary conditions. This illustrates the profound impact boundary conditions may have on solutions. This third argument highlights perhaps further the fact that the first two arguments do not need to be quantum mechanical and could be entirely classical. What would not be classical is Bohr’s overinterpretation, which we can now appreciate to be wrong.
- The three arguments given up to now recollect what we already developed in [2]. A fourth argument consists in referring to Gleason’s theorem which is obviously pure mathematics (and also does not depend on the Born rule). But this is of course very similar to the previous arguments. We list these four arguments as different arguments only because they might look different at first sight. When there is a hard nut to be cracked in order to solve a mathematical problem, one will forcedly hit it whatever the road one takes in trying to solve the problem. Call it a conservation law for hard nuts. By changing the approach we may only discover the hard nut in a different guise.

---

<sup>2</sup> A spinor in  $SU(2)$  is a rotation [5], not a vector. Vector decompositions of spinors and rotations are therefore meaningless.

• In summary, we have been aware of this kind of objections for a long time, but some way they have been overlooked in deriving the inequalities, perhaps because it was considered that the objections were quantum mechanical and not classical such that one should not consider them in something that was supposed to represent classical thinking as opposed to quantum mechanical thinking! But what is here associated with classical thinking is only poor mathematical thinking based on “physical intuition” applied to problems that are purely mathematical! What is wrong and pollutes the whole debate are the overinterpretations of the mathematics the Copenhagen interpretation is teeming with and are supposed to define “quantum intuition”. We must become aware of the fact that the Copenhagen interpretation has brainwashed us with the agitprop that mathematically wrong intuition would be a prerogative of classical rationalism while “quantum intuition” would be exact.

## 4 Would-be simple scenarios for obtaining a common probability distribution that fail

### 4.1 A logically correct extension based on unions of sets

Let us now explain why enlarging the sets  $V_{jk}$  to a set  $V$  in such a way that  $p$  will engulf all probabilities  $p_{jk}$  is all but trivial. It is obviously already in contradiction with the four arguments given above. In trying to follow our intuition and to define a common distribution function  $p$  one will run into all sorts of difficulties, which is normal because they are there to prevent us from deriving a contradiction from the mathematics. But as physicists we have been taught to take our strides with mathematical rigor, such that we are prone to make some booby traps go off. Very often we get away with our lack of rigor, but not this time. This time we have paid very dearly.

Let us try to define an extension.<sup>3</sup> In the following  $m$  stands actually for a compound index of the type  $(jk)$  above. We note the sets  $V_{jk}$  as  $V_m$ . We note an element  $\lambda \in V_{jk}$  as  $x_m \in V_m$ , and we note the probability distribution over  $V_m$  as  $p_m \in F(V_m, [0, \infty[)$ . We have thus  $\int_{V_m} p_m(x_m) dx_m = 1$ . We want to make an extension that includes two sets of the type  $V_m$ , which we will call  $V_1$  and  $V_2$ . In reality, we will need four of them because we have four different set-ups. But the example with two sets will show the method works in general. The reader might think that our approach is awkward and object that there is a much more elegant one that does not lead to a normalization problem. But we will see that this elegant approach is logically flawed. The present approach is not flawed. The only problem it has will turn out to be its normalization problem. We define:

$$\bar{V}_m = \{(x_m, m) \mid x_m \in V_m\}, \quad V = \bar{V}_1 \cup \bar{V}_2, \quad (9)$$

We have then:

$$\bar{V}_1 \cap \bar{V}_2 = \emptyset, \quad (10)$$

such that we avoid the problem that  $V_1$  and  $V_2$  can physically spoken not exist simultaneously because two physical objects like polarizers cannot be present at a same place at the same time. Since we have:

$$\forall(x, \ell) \in V : \begin{cases} \ell = 1 \Rightarrow \exists x_1 \in V_1 \mid x = x_1 \\ \ell = 2 \Rightarrow \exists x_2 \in V_2 \mid x = x_2 \end{cases} \quad (11)$$

we can define:

$$\forall(x, \ell) \in V : \begin{cases} \ell = 1 \Rightarrow p(x, \ell) = p_1(x_1) \\ \ell = 2 \Rightarrow p(x, \ell) = p_2(x_2) \end{cases} \quad (12)$$

We can then write:

$$p_m(x_m) = p(x, \ell) \delta_{m\ell} \quad (13)$$

That would permit to write the two probability distributions as one distribution. But this leads to a normalization problem, because:

$$\int_V p(x, \ell) dx = \sum_{m=1}^2 \int_{V_m} p_m(x_m) dx_j = 2. \quad (14)$$

<sup>3</sup> Defining in an alternative approach everything in terms of  $V = V_{11}$ , will keep us in the presence of 4 different distributions  $p_{jk}$ .

We would thus have to allow for a renormalization. An extension  $V_{jk} \rightarrow V$  changes the normalizations of the probabilities  $p_{jk}$ . In the approach that treats the four set-ups, the quantities  $\frac{1}{2} \cos^2(\alpha_j - \beta_k)$  would then have to be renormalized to  $\frac{1}{8} \cos^2(\alpha_j - \beta_k)$  in applying the inequality to the theoretical data.

## 4.2 A logically flawed extension based on Cartesian products of sets

Let us explore now the more elegant scenario. Let us write it out in full for the four possible joint polarizer settings. We put  $V_{jk} = A_j \times B_k \times P$ . Here  $A_j$  and  $B_k$  are the sets of hidden variables in the two polarizers.  $P$  is the set of hidden variables of the photons. Every set-up contains the hidden variables  $A_j$ ,  $B_k$ , and  $P$ . To obtain a common description we must use  $A_1 \times A_2 \times B_1 \times B_2 \times P$ . Let us note the elements of  $A_j$  by  $\lambda_j$ , those of  $B_k$  by  $\mu_k$ , and those of  $P$  by  $\zeta$ . The function  $p_{jk}$  is defined over  $A_j \times B_k \times P$ . Suppose that  $p_{jk}(\lambda_j, \mu_k, \zeta)$  can be written as a product  $f_j(\lambda_j)g_k(\mu_k)h(\zeta)$ . In fact, the properties of a polarizer do not depend on the properties of the photons and *vice versa*. The properties of a polarizer do also not depend on the properties of another polarizer. Therefore the distributions over the three sets must be independent, such that we can multiply them. We now also need the characteristic functions of the sets  $A_j$  and  $B_k$ . Call  $\bar{a}_j \in F(A_j, S)$  the characteristic function of  $A_j$  defined by:  $\forall \lambda_j \in A_j \Rightarrow \bar{a}_j(\lambda_j) = 1$ ,  $\forall \lambda_j \notin A_j \Rightarrow \bar{a}_j(\lambda_j) = 0$ . Analogously, we call *mutatis mutandis*  $\bar{b}_k \in F(B_k, S)$  the characteristic function of  $B_k$  defined by:  $\forall \mu_k \in B_k \Rightarrow \bar{b}_k(\mu_k) = 1$ ,  $\forall \mu_k \notin B_k \Rightarrow \bar{b}_k(\mu_k) = 0$ .

Let us now write the term that contains  $a_1 b_1$  in this formalism. As  $a_1 b_1 = a_1 b_1 \bar{a}_2 \bar{b}_2$ , because the added factors are equal to 1, we have:

$$\begin{aligned} \int_{A_1 \times A_2 \times B_1 \times B_2 \times P} a_1(\lambda_1, \zeta) \bar{a}_2(\lambda_2) b_1(\mu_1, \zeta) \bar{b}_2(\mu_2) \cdot f_1(\lambda_1) f_2(\lambda_2) g_1(\mu_1) g_2(\mu_2) h(\zeta) \cdot d\lambda_1 d\lambda_2 d\mu_1 d\mu_2 d\zeta = \\ \int_{A_1 \times B_1 \times P} a_1(\lambda_1, \zeta) b_1(\mu_1, \zeta) f_1(\lambda_1) g_1(\mu_1) h(\zeta) d\lambda_1 d\mu_1 d\zeta \int_{A_2} \bar{a}_2(\lambda_2) f_2(\lambda_2) d\lambda_2 \int_{B_2} \bar{b}_2(\mu_2) g_2(\mu_2) d\mu_2 = \\ \int_{A_1 \times B_1 \times P} a_1(\lambda_1, \zeta) b_1(\mu_1, \zeta) f_1(\lambda_1) g_1(\mu_1) h(\zeta) d\lambda_1 d\mu_1 d\zeta \int_{A_2} f_2(\lambda_2) d\lambda_2 \int_{B_2} g_2(\mu_2) d\mu_2 = \\ \int_{A_1 \times B_1 \times P} a_1(\lambda_1, \zeta) b_1(\mu_1, \zeta) f_1(\lambda_1) g_1(\mu_1) h(\zeta) d\lambda_1 d\mu_1 d\zeta. \quad (15) \end{aligned}$$

As we can do this for every term in the Bell inequality, this seems to suggest that  $f_1 f_2 g_1 g_2 h \in F(A_1 \times A_2 \times B_1 \times B_2 \times P, [0, \infty])$  could be the common distribution desired. This would then be the existence proof. But, once again, this proof is flawed. In fact, the photons do not explore the full set  $A_1 \times B_1$ . They only explore a subset of  $A_1 \times B_1$ . E.g. if  $C_1 \subset \mathbb{R}^3$  is the space occupied by the set of molecules that intervene in  $A_1$ , and  $D_1 \subset \mathbb{R}^3$  the space occupied by the set of molecules that intervene in  $B_1$ , then the photons do not explore  $C_1 \times D_1$ . If  $\mathbf{r}_1 \in C_1$  is explored by a photon, then only  $-\mathbf{r}_1 \in D_1$  will be explored by its buddy photon, not the full set  $D_1$ . The set of coordinates explored is thus a set of correlated pairs  $(\mathbf{r}_1, -\mathbf{r}_1)$ , rather than the set  $C_1 \times D_1$  of uncorrelated pairs  $(\mathbf{r}_1, \mathbf{r}_2)$ . Hence, although the probability distributions in the two polarizers are independent, their definition domains are correlated because the photons are correlating them! We can thus not factorize the sets as we have done. We cannot write  $V_{jk} = A_j \times B_k \times P$ , and therefore there is no common set  $A_1 \times A_2 \times B_1 \times B_2 \times P$  on which we could describe simultaneously all different terms that occur in the inequality. Despite the intuitive appeal and *prima facie* success of the second method, the first method is thus actually logically superior, because we refrain from trying to decompose  $V_{jk}$  and this way we avoid making the error that occurs in the second method.

## 4.3 Caveats

Consider a first person who rebuts the first repair scenario. He might be unaware of the second repair scenario and think then that he has proved conclusively that the Bell inequalities are wrong, while he might be considered as a fool by a second person who believes in the second repair scenario and is not aware of its errors and of the first repair scenario. He will believe that the first person is stupid and completely has missed the point. But the situation is completely symmetrical. The first person may consider the second person as foolish. In reality both persons are trying to contribute to science in a very serious constructive way. They are not at all ridiculous, but they are both falling prey to unfair scorn because the incompleteness of the proof of the Bell inequality turns the context into a minefield. Clauser *et al.* should not have introduced a tacit assumption without proof. They should have given a complete proof without any possible loopholes. Instead of that people are now kept guessing if a complete proof could exist and on what kind of approach it could be based. A correct proof must only leave it to people to check if it is true, and not

leave it to people to fill some gaps. The proof that has been given is wrong because it is at the very best incomplete. It forces people then to imagine a plethora of possible approaches that could be attempted in order complete it, because they would like to prove that the inequality is wrong, after figuring out that the proof is wrong.

In fact, as long as we have not proved that the inequality is wrong, we permanently risk that people may come up with alternative scenarios whereby a subtle point has been missed. They will then unwittingly revert the charge of proof, as it becomes then the responsibility of others to figure out in great pains where the unsuspected error occurs. We mentioned one alternative and its error in Footnote 3. It is for this reason that it cannot possibly be our task to prove that Eq. 8 is wrong. This reversal of the charge of proof has already been lingering on for several decades now. In the two examples given above, the error was hiding in a little corner. In both cases I went through immense frustration because for a long time I just could not spot the error that would come to my rescue. That was also the case with the tacit *Ansatz* of a common probability distribution in the first place. In the situation as it stands, the inequality has just not been proved.

Nieuwenhuizen [6] has also pointed out how the tacit assumption of a common probability distribution ruins the Bell inequalities, actually a long time before me. I was not aware of his work and figured this out independently. I hope that at least my work may strengthen the case, because entanglement would have been a very serious and puzzling conceptual issue. It is important to point out that in our approach the angle  $\alpha_j - \beta_k$  is *non-locally defined*, without any need for signalling and without any violation of relativity, such that this non-locality is not an issue [2]!

#### 4.4 Epilogue

We noted that we could have stopped our rebuttal of the Bell inequalities at the end of Section 2. All the additions are just forced upon us by physicists, who are inert to requests for extreme rigor, such that we are forced to anticipate that they could pooh-pooh the objection in Section 2 despite the fact that it is a completely pertinent, terrible blow to the derivation of the Bell inequalities. We do not have the charge of proof to show that Eq. 8 is wrong. It is mathematically clear enough that the necessity to provide an existence proof is absolutely compelling. Asking more from me amounts to reverting the charge of proof. It cannot be anybody's task to imagine all possible scenarios to define a common probability distribution and prove them wrong. We only have to express reasonable doubt about the hidden assumption, making the reader wonder if it really goes without saying that we can introduce it without proof, under the pretext that it would be self-evident.

The discussion shows that the derivation of the Bell inequality is not at all as obvious and trite as we were inclined to believe and that the purely mathematical problem of the existence of a common probability distribution is indeed the pivotal point on which the whole Bohr-Einstein debate hinges. In the derivation of the Bell inequalities, this issue has been swept unintentionally under the carpet. The lesson we must learn from this is that homespun intuition is bound to fail beyond a certain level of mathematical subtlety. Especially in probability calculus, which is fraught with nasty pitfalls, the danger of making an error is permanently lurking. This also transpires from our analysis of the double-slit experiment [2]. One should therefore eschew tackling deep foundational questions with a strategy that crucially depends on probability calculus.

### 5 Appendix: The quadratic character of the Born rule

We may think that the Born rule is eminently non-classical. But the quadratic character of the Born rule is an unavoidable consequence of the fact that vectors and four-vectors are “quadratic” rank-2 expressions in terms of spinors in the Lorentz group and in the rotation group [5]. It is just group theory. Let us give the reader some feeling for this idea. For the proof that the unit vectors  $\mathbf{e}_x$ ,  $\mathbf{e}_y$  and  $\mathbf{e}_z$  of  $\mathbb{R}^3$  are quadratic expressions of the spinors of the rotation group  $SU(2)$ , we refer the reader to reference [5], more specifically to Eqs. 30-32. These equations tell us that for the basis vectors  $\mathbf{e}_x = (x_1, y_1, z_1)$ ,  $\mathbf{e}_y = (x_2, y_2, z_2)$ ,  $\mathbf{e}_z = (x_3, y_3, z_3)$  of a rotated frame we have:

$$\begin{aligned} x_1 &= \frac{1}{2}(\xi_0^2 - \xi_1^2 + \xi_0^{*2} - \xi_1^{*2}), & y_1 &= \frac{i}{2}(\xi_0^2 + \xi_1^2 - \xi_0^{*2} - \xi_1^{*2}), & z_1 &= -(\xi_0\xi_1 + \xi_0^*\xi_1^*), \\ x_2 &= \frac{i}{2}(-\xi_0^2 + \xi_1^2 + \xi_0^{*2} - \xi_1^{*2}), & y_2 &= \frac{1}{2}(\xi_0^2 + \xi_1^2 + \xi_0^{*2} + \xi_1^{*2}), & z_2 &= (\xi_0\xi_1 - \xi_0^*\xi_1^*), \\ x_3 &= \xi_0\xi_1^* + \xi_0^*\xi_1, & y_3 &= i(\xi_0\xi_1^* - \xi_0^*\xi_1), & z_3 &= \xi_0\xi_0^* - \xi_1\xi_1^*. \end{aligned} \quad (16)$$

Here  $\xi_0$  and  $\xi_1$  are the components of the spinor  $\boldsymbol{\xi}$ :

$$\boldsymbol{\xi} = \begin{bmatrix} \xi_0 \\ \xi_1 \end{bmatrix}, \text{ which serves as a shorthand for the } SU(2) \text{ matrix: } \mathbf{R}_0 = \begin{bmatrix} \xi_0 & -\xi_1^* \\ \xi_1 & \xi_0^* \end{bmatrix}, \quad (17)$$



of the rotation that was applied to the pristine canonical frame to obtain  $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ . By covariance, any vector  $\mathbf{r} \in \mathbb{R}^3$  is a quadratic expression in terms of spinors. To generalize Eq. 16 to vectors  $\mathbf{r}$  that are not unit vectors, we must use generalized spinors  $\psi = \sqrt{r} \xi$ . This can be generalized to Minkowski space-time. The four basis vectors of Minkowski space time are quadratic expressions of Lorentz spinors, and four-vectors in general quadratic expressions of spinors. The reader can get a feeling for this generalization from reference [5] where we generalize the results for the rotation group  $SU(2)$  in  $\mathbb{R}^3$  to rotation groups in  $\mathbb{R}^n$ .

We know from relativity that charge and current density  $\rho$  and  $\mathbf{j}$  are part of a charge-current density four-vector  $(c\rho, \mathbf{j})$  for electric charges. A probability density is part of an analogous four-vector for “probability charges”. We can call this the probability charge-current density four-vector. To stress the analogy, we will write this four-vector in what follows also as  $(c\rho, \mathbf{j})$ , such that we will use both  $\rho$  and  $p$  to refer to the probability density.

In the representation  $SL(2, \mathbb{C})$  of the Lorentz group a four-vector  $(c\rho, \mathbf{j})$  is expressed as  $c\rho \mathbb{1} + \mathbf{j} \cdot \boldsymbol{\sigma}$  (see e.g. reference [1], Chapter 4). It transforms under Lorentz transformations  $\mathbf{L} \in SL(2, \mathbb{C})$  according to:

$$c\rho \mathbb{1} + \mathbf{j} \cdot \boldsymbol{\sigma} = \begin{bmatrix} c\rho + j_z & j_x - \imath j_y \\ j_x + \imath j_y & c\rho - j_z \end{bmatrix} \rightarrow \mathbf{L} \begin{bmatrix} c\rho + j_z & j_x - \imath j_y \\ j_x + \imath j_y & c\rho - j_z \end{bmatrix} \mathbf{L}^\dagger \quad (18)$$

Four-vectors are bilinear covariants, i.e. quadratic expressions in terms of spinors  $\xi$ , which transform linearly:  $\xi \rightarrow \mathbf{L}\xi$  under Lorentz transformations. The quadratic transformation law Eq. 18 is just another manifestation of the fact that four-vectors are quadratic expressions of spinors. One therefore often states that a spinor is the “square root of a vector”.

The ultimate non-relativistic limit corresponds to  $\mathbf{v} = \mathbf{0}$ . We have then  $\mathbf{j} = \mathbf{0}$  and  $\mathbf{L}$  becomes a rotation  $\mathbf{R}_0$ , as the Lorentz group reduces to the rotation group  $SU(2)$ , which is embedded within  $SL(2, \mathbb{C})$ . The four-vector reduces then to a scalar, because it becomes  $(c\rho, \mathbf{0})$ . For probability charges this scalar has to be positive-definite and of degree 2 in terms of a spinor  $\xi$ , because, as said, four-vectors are quadratic expressions of spinors. The most simple quadratic expression for such a positive-definite scalar that comes to one’s mind is  $p[\xi^\dagger \xi]$ . We have then  $\chi = \sqrt{p} \xi$ , and  $p = \chi^\dagger \chi$ . Well, to be quite honest, the solution is not unique. We could actually also propose  $\chi^\dagger \mathbf{L}^\dagger \mathbf{L} \chi$ , for some  $\mathbf{L} \in SL(2, \mathbb{C})$ , but this would only correspond to a change of basis:  $\chi \rightarrow \mathbf{L}\chi$  by a group transformation. The probability density is thus a quadratic expression in terms of a spinor field  $\chi$ , that takes the values  $\chi(\mathbf{r}, t)$ . What we still would have to explain is why this spinor field  $\chi$  is actually the wave function  $\psi$ .

As explained, we abandoned the idea of explaining this in the present paper because the proof becomes really too involved. We can just give a sketch of the idea behind it. The spinor  $\xi$  represents a group element. It corresponds thus to a Lorentz transformation such that we can identify it with a Lorentz frame. For a single frame we have  $\xi^\dagger \xi = 1$  (or  $\xi^\dagger \xi = \gamma$  relativistically). This is the essence of what we explained above. Hence when we put  $\psi = \sqrt{p} \xi$ ,  $p$  is actually a frame density. At this point we start needing some physics such that the argument will no longer be pure group theory, but the reasoning will be purely classical. In constructing the wave equation we attach a frame to each electron because the frame describes its dynamical state in terms of an instantaneous Lorentz transformation. Therefore the frame density  $p$  is the electron density, i.e. the probability density.

Of course, the detailed argument is quite elaborate because it relies on following it throughout the derivation of the Dirac equation given in [1]. We do not want all this to interfere with our discussion of the Bell inequalities. Our first two arguments of the discussion in Section 3 remain valid, even if we consider the Born rule as non-classical. It is only that a proof that the Born rule is classical would render these arguments even more poignant. It would prove that Eq. 8 is even wrong in classical mechanics and that our arguments do not rely on smuggling in “quantum effects”. But the other two arguments are already pointing in the same direction. At least it must be clear from this section that the strangest aspect of the Born rule, viz. its quadratic character, is just a consequence of group theory. The probability density is thus a quadratic expression in terms of a spinor field  $\chi$ , that takes the values  $\chi(\mathbf{r}, t)$ , whereby it takes further physical arguments to prove that  $\chi = \psi$ .

The derivation above is valid within the context of the Schrödinger equation, where we neglect the relativistic effects of boosts. In a moving frame, we would have  $\rho \rightarrow \gamma\rho$ , but as  $\gamma \approx 1$ , its expression will not change in a non-relativistic context. We will only have also a current density  $\mathbf{j} \neq \mathbf{0}$  when  $\mathbf{v} \neq \mathbf{0}$ . In the Schrödinger equation we actually only treat the phase of  $\psi$ , considering the other components to be constant, such that we can drop them tacitly from the formalism. In other words, we replace:

$$\xi = e^{\imath\varphi} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightarrow \xi = e^{\imath\varphi}, \quad \xi^\dagger \xi \rightarrow \xi^* \xi. \quad (19)$$

such that  $\psi \rightarrow \psi = \sqrt{p} e^{\imath\varphi}$  with  $p = \psi^* \psi$ . Within the Dirac theory, the definitions must be adjusted, but the expressions always follow from group theory. Strictly spoken, all this applies only to particles with spin 1/2. The Schrödinger equation is also used for other particles endowed with mass like He nuclei or atoms, which are described by other representations of the Lorentz group and the rotation group, such that these cases would require other proofs.

## References

1. G. Coddens, in *From Spinors to Quantum Mechanics*, Imperial College Press, London (2015).
2. G. Coddens, <https://hal-cea.archives-ouvertes.fr/cea-01383609v4>.
3. G. Coddens, <https://arxiv.org/pdf/1309.6057.pdf>. This paper contains an error.
4. A. Shimony in *The New Physics*, Paul Davies ed., Cambridge University Press, Cambridge (1983), pp. 373-395.
5. An introduction to  $SU(2)$  and  $SO(3)$  is given in <https://hal-cea.archives-ouvertes.fr/cea-01572342>, combined with Chapter 3 of [1]. One cannot address quantum mechanics without a perfect mastery of spinors and group theory.
6. T.M. Nieuwenhuizen, *Found. Phys.* (2011) **41**, 580-591.