Examples of mathematical beauty when comparing classical and quantum worlds

Adán Cabello

Departamento de Física Aplicada II, Universidad de Sevilla, 41012 Sevilla, Spain
adan@us.es

1 The golden ratio in Hardy’s proof of Bell’s theorem

The following is my favorite of all the anecdotes involving Alberto Galindo I know. I was preparing my Ph. D. in the Theoretical Physics Department of the Universidad Complutense of Madrid. One day, my supervisor, Guillermo García Alcaine, and I went to Galindo’s office and left one of our manuscripts for him to review. Basically, the paper concerned the following problem: given two distant parties sharing two qubits, $A$ and $B$, prepared in an entangled state $\eta$ such that

\begin{align}
P_\eta(A_1 = 1, B_1 = 1) &= \mathcal{P} > 0, \quad (1) \\
P_\eta(A_0 = 1, B_1 = 1) &= 0, \quad (2) \\
P_\eta(A_1 = 1, B_0 = 1) &= 0, \quad (3) \\
P_\eta(A_0 = -1, B_0 = -1) &= 0, \quad (4)
\end{align}

where $A_0$ and $A_1$ ($B_0$ and $B_1$) are two alternative dichotomic experiments (i.e., having only two possible outcomes, which we label $\pm 1$) on qubit $A$ ($B$), which is the maximum possible value for $\mathcal{P}$? We waited for some days and then went back to Galindo’s office and asked what he thought of it. “I haven’t had time to read it,” he said. We felt a bit disappointed, but after a few seconds of embarrassing silence, Galindo added “However, your result is correct.” Guillermo and I looked incredulously at each other. After a few more seconds, we gathered the courage to ask him what were his reasons for thinking that. He just said “It’s correct because the final result is the fifth power of the inverse of the golden ratio, and such a beautiful number cannot happen by chance.” Of course, he was right; our result was

\[ \mathcal{P}_{\text{max}} = \left( \frac{\sqrt{5} - 1}{2} \right)^5. \quad (5) \]

Unfortunately, we later found out that such a beautiful result was previously discovered by Lucien Hardy [1], as part of what, according to David Mermin, is
"the best version of Bell's theorem" [2], which is based on equations (1), (2), (3) and (4).

Bell's theorem [3] points out one of the main differences between classical and quantum mechanics; it states that it is impossible to "complete" quantum mechanics in such a way that local observables whose results can be predicted with certainty from space-like separated measurements have predefined values "revealed" by the experiments, as in classical mechanics. We will not review Hardy's proof here; for details, see [1, 2]. However, in section 2 we shall present another surprising example of how irrational numbers appear when we compare classical and quantum physics, and in section 3 we shall present another example of "mathematical beauty", namely, the simplest possible proof of another fundamental result, the Kochen-Specker theorem [4, 5, 6].

2 Pi in the sky\(^1\) between classical and quantum worlds

2.1 Introduction

"Pour moi il n'y a pas de Dieu, il y a π."

H. Cartier-Bresson [7].

"For me, God does not exist, it only exists π," said Henri Cartier-Bresson, the famous photographer, a year before his death. Through the years, I have learned that irrational numbers appears in the most unexpected and diverse contexts. Particularly, π appears in so many different situations that there are very thick tomes devoted to this (for instance, see [8, 9]). Therefore, it would be dishonest to show genuine surprise when finding π once again, even in the most unexpected circumstances. What is remarkable here, at least to me, is that, to my knowledge, the following simple but somehow fundamental calculations cannot be found elsewhere in the literature [10].

2.2 How larger quantum correlations are than classical ones?

Quantum information (that is, information carried by microscopic systems described by quantum mechanics such as atoms or photons) can connect two space-like separated observers by correlations that cannot be explained by classical communication. This fact, revealed by Bell's inequalities and violations thereof [3, 11], is behind common statements such as that quantum correlations are "stronger" or "larger" than classical ones, or that quantum-mechanical systems may be "further correlated" than those obeying classical physics (for instance, see [12]), and has

\(^1\)"Pi in the sky" sounds like "pie in the sky", which is a common expression for saying "castle in the air"; "Pi in the sky" is also the title of a short science-fiction story by Fredric Brown by around 1945.
been described as "the most profound discovery of science" [13]. Given its fundamental importance, it is surprising that the question of how "larger" than classical correlations quantum correlations are has not, to my knowledge, a precise answer beyond the fact that quantum mechanics violates Clauser-Horne-Shimony-Holt (CHSH) inequalities [11] up to \( 2\sqrt{2} \) (Tsirelson's bound [14]), while the classical bound is just \( 2 \) [11]. To be more specific, if we denote by \( \mathcal{Q} \) the set of all correlation functions allowed by quantum mechanics in a given experimental scenario, and by \( \mathcal{C} \) the corresponding set of correlation functions allowed by a general classical deterministic theory, a more precise measure of how larger quantum correlations are compared to classical ones would be the ratio between the volumes (i.e., hyper-volumes or contents) of both sets, \( V_{\mathcal{Q}}/V_{\mathcal{C}} \).

Another interesting problem is why quantum correlations cannot be even "larger" than they are [15]. How much "larger" could, in principle, the set of correlations be? A precise measurement of this would be the ratio between the volume of quantum correlations and the volume of the set \( \mathcal{L} \) of all possible correlations allowed by any general probabilistic local causal theory, \( V_{\mathcal{Q}}/V_{\mathcal{L}} \).

The EPR-Bell scenario [3, 11, 16, 17] is the simplest and most basic one where the difference between classical and quantum correlations arises. It consists of two alternative dichotomic experiments (i.e., having only two possible outcomes, which we can label \( \pm 1 \)), \( A_0 \) or \( A_1 \), on a subsystem \( A \), and other two alternative dichotomic experiments, \( B_0 \) or \( B_1 \), on a distant subsystem \( B \). Therefore, for the EPR-Bell scenario the set of correlations \( \langle A_i B_j \rangle \) is 4-dimensional. On the other hand, this is the most basic scenario, since it is contained in any other experimental scenario involving more subsystems, more experiments per subsystem, or more outcomes (a discrete number of them) per experiment.

### 2.3 Correlations allowed by a classical deterministic theory

Froissart [18] and Fine [19, 20] (see also [21, 22]) proved that, for the EPR-Bell scenario, the set of all joint probabilities attainable by a classical deterministic local theory (i.e., a theory in which the local variables of a subsystem determine the results of local experiments on this subsystem) is an 8-dimensional polytope with 16 vertices and 24 faces. The set of correlations is a 4-dimensional projection of the set of joint probabilities. The connection between both sets is given by

\[
\langle A_i B_j \rangle = \sum_{k,l \in \{-1,1\}} a_k b_l P(A_i = a_k, B_j = b_l).
\]  

(6)

The 4-dimensional projection corresponding to the set \( \mathcal{C} \) of all correlation functions that can be attained by a classical deterministic local theory is defined by 8 CHSH inequalities. To be precise, a set of 4 real numbers \( \langle A_i B_j \rangle \) \((i, j = 0, 1)\) belongs to \( \mathcal{C} \), i.e., represents a set of correlations attainable by a classical deterministic local theory, if and only if

\[
|\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle + \langle A_1 B_1 \rangle - 2\langle A_i B_j \rangle| \leq 2,
\]  

(7)
for all \( i, j = 0, 1 \). The volume of this four-dimensional set \( \mathcal{C} \) can be easily calculated,

\[
V_\mathcal{C} = \frac{2^5}{3}.
\]  

(8)

2.4 Correlations allowed by a general probabilistic local theory

Supposing we do not assume that the results of local experiments on the subsystems are determined, as in the previous case, but are probabilistic. The only restriction now will be that signaling is forbidden (i.e., the two distant observers cannot signal to one another via their choice of input). The no-signaling condition restricts the set of joint probabilities. The no-signaling condition imposes that the marginal probabilities \( P(B_j = b) \) \( P(A_i = a) \) should be independent of the choice of \( A_i \) \( B_j \), for all \( B_j \) and \( b \in \{-1, 1\} \) [for all \( A_i \) and \( a \in \{-1, 1\} \)]. This implies 8 restrictions on the set of joint probabilities, so that the set of all possible joint probabilities satisfying the no-signaling condition has dimension 8. This set is a convex polytope with 24 vertices and 16 faces [23]. However, the restrictions imposed on the set of joint probabilities by the no-signaling condition do not imply new non-trivial restrictions on the set of correlations \( \langle A_i B_j \rangle \). There are either sets of joint probabilities violating no-signaling but satisfying inequalities (7) and sets satisfying no-signaling but maximally violating (7) [15]. Therefore, the set \( \mathcal{C} \) of all correlation functions that can be attained by a probabilistic local theory is simply defined by the 8 inequalities

\[
|\langle A_i B_j \rangle| \leq 1,
\]  

(9)

for \( i, j = 0, 1 \). \( \mathcal{C} \) is a 4-dimensional cube (a tesseract). Its volume is

\[
V_\mathcal{C} = 2^4.
\]  

(10)

Comparing (8) and (10), it is easy to see that the volume of the set of correlations attainable by classical deterministic theories is just \( 2/3 \) of that allowed by probabilistic local theories.

2.5 Correlations allowed by quantum mechanics

Although rarely mentioned in the literature, to my knowledge, there are three equivalent sets of necessary and sufficient conditions to define the set \( \mathcal{Q} \) of correlations attainable by quantum mechanics. The first was provided by Tsirelson [24]. According to Tsirelson, a set of 4 correlations \( \langle A_i B_j \rangle \) \( (i, j = 0, 1) \) is realizable in quantum mechanics (i.e., belongs to \( \mathcal{Q} \)) if at least one of the following two inequalities holds:
0 \leq \left( \langle A_0 B_1 \rangle \langle A_1 B_0 \rangle - \langle A_0 B_0 \rangle \langle A_1 B_1 \rangle \right) \\
\times \left( \langle A_0 B_0 \rangle \langle A_0 B_1 \rangle - \langle A_1 B_0 \rangle \langle A_1 B_1 \rangle \right) \\
\times \left( \langle A_0 B_0 \rangle \langle A_1 B_0 \rangle - \langle A_0 B_1 \rangle \langle A_1 B_1 \rangle \right) \\
\leq \frac{1}{4} \left( \sum_{i,j} \langle A_i B_j \rangle^2 \right)^2 - \frac{1}{2} \sum_{i,j} \langle A_i B_j \rangle^4 - 2 \prod_{i,j} \langle A_i B_j \rangle, \quad (11)

0 \leq 2 \max_{i,j} \langle A_i B_j \rangle^4 - (\max_{i,j} \langle A_i B_j \rangle^2) \left( \sum_{i,j} \langle A_i B_j \rangle^2 \right) \\
+ 2 \prod_{i,j} \langle A_i B_j \rangle. \quad (12)

The second characterization of $Q$ is due to Landau [25]. According to him, 4 correlations belong to $Q$ if and only if they satisfy the following inequalities:

$$|\langle A_0 B_0 \rangle \langle A_0 B_1 \rangle - \langle A_1 B_0 \rangle \langle A_1 B_1 \rangle| \leq$$
$$\sqrt{1 - \langle A_0 B_0 \rangle^2} \sqrt{1 - \langle A_0 B_1 \rangle^2} + \sqrt{1 - \langle A_1 B_0 \rangle^2} \sqrt{1 - \langle A_1 B_1 \rangle^2}. \quad (13)$$

These inequalities (13) are equivalent to inequalities (11) and (12) [23].

The third equivalent definition of $Q$ can be explicitly found for the first time in [23] (although it can be easily derived from the results in [25]). According to this, 4 correlations belong to $Q$ if and only if they satisfy the following 8 inequalities:

$$| \arcsin \langle A_0 B_0 \rangle + \arcsin \langle A_0 B_1 \rangle + \arcsin \langle A_1 B_0 \rangle + \arcsin \langle A_1 B_1 \rangle - 2 \arcsin \langle A_i B_j \rangle | \leq \pi, \quad (14)$$

for all $i, j = 0, 1$. Using inequalities (14) to describe $Q$ has the advantage of being analogous to using inequalities (7) to describe $C$. These inequalities (14) have been recently rediscovered by Masanes [26].

The simplest way to calculate the volume of $Q$, which is a 4-dimensional convex set [23], is by using expression (13). Then, it can be seen that

$$V_Q = \frac{3\pi^2}{2} \approx 0.925 \times 2^4. \quad (15)$$

Therefore, the ratio between the volumes of the set of quantum correlations and those allowed by classical deterministic theories, which is a good measure of how larger than classical correlations quantum correlations are for the EPR-Bell scenario, is

$$\frac{V_Q}{V_C} = \left( \frac{3\pi}{8} \right)^2 \approx 1.388. \quad (16)$$

Is not surprising to find $\pi$ in the answer to the question of how larger than classical correlations quantum correlations are?

On the other hand, the ratio between the volumes of the set of quantum correlations and those allowed by general local theories theories is
\[
\frac{V_Q}{V_C} = \frac{3\pi^2}{32} \approx 0.925.
\] (17)

This result allows us to quantify how "larger" than the set of quantum correlations the set of possible correlations could be: 7.5% of the, in principle, possible sets of four correlations never occur in nature.

3 Quantum coffee, Wheeler and the simplest proof of the Kochen-Specker theorem

3.1 Wheeler's "it from bit"

"Trying to wrap my brain around this idea of information theory as the basis of existence, I came up with the phrase "it from bit." The universe in all that it contains ("it") may arise from the myriad yes-no choices of measurement (the "bits"). Niels Bohr wrestled for most of his life with the question of how acts of measurement (or "registration") may affect reality. It is registration (...) that changes potentiality into actuality. I build only a little on the structure of Bohr's thinking when I suggest that we may never understand this strange thing, the quantum, until we understand how information may underlie reality. Information may not be just what we learn about the world. It may be what makes the world. An example of the idea of it from bit: When a photon is absorbed, and thereby "measured"—until its absorption, it had no true reality—an unsplittable bit of information is added to what we know about the world, and, at the same time that bit of information determines the structure of one small part of the world. It creates the reality of the time and place of that photon's interaction."

J. A. Wheeler [and K. W. Ford] [27], pp. 340-341.

The first time I heard about John Wheeler's way of understanding quantum mechanics (see also [28, 29]) and thought, and I still do, that it should be the "correct" point of view, was during an informal chat that Alberto Galindo shared with us, his then students of fourth curse of the degree in Physics, over a cup of coffee. The idea that the measuring process creates a "reality" that did not exist objectively before the intervention since, as Asher Peres likes to say, "unperformed experiments have no results" [30], is supported by the Kochen-Specker (KS) theorem [4, 5, 6], one of the most fundamental results in quantum mechanics. Alberto Galindo was also presiding the board of examiners of my Ph. D. thesis [31]. One of the main results of that thesis was a proof [32, 33, 34], the simplest known at the time, of the KS theorem. This proof, which led to many interesting new results, is reviewed below. With the exception of the figure illustrating the proof, there is nothing new in the subsections that follow. The pretext for presenting it
again is not only because such a proof still is the simplest one known, but also because, as has been recently proved [35, 36], it is the simplest possible proof of the KS theorem, as conjectured by Peres [37].

3.2 The Kochen-Specker theorem

The KS theorem states that yes-no questions about an individual physical system cannot be assigned a unique answer in such a way that the result of measuring any mutually commuting subset of these yes-no questions can be interpreted as revealing these preexisting answers. More precisely, the KS theorem asserts that, in a Hilbert space $\mathcal{H}_d$ with a finite dimension, $d \geq 3$, it is possible to construct a set of $n$ projection operators, which represent yes-no questions about an individual physical system, so that none of the $2^n$ possible sets of “yes” or “no” answers is compatible with the sum rule of quantum mechanics for orthogonal resolutions of the identity (i.e., if the sum of a subset of mutually orthogonal projection operators is the identity, one and only one of the corresponding answers ought to be “yes”). This conclusion holds irrespective of the quantum state of the system. Implicit in the KS theorem is the assumption of noncontextuality: each yes-no question is assigned a single unique answer, independent of which subset of mutually commuting projection operators one might consider it with. Therefore, the KS theorem discards hidden-variable theories with this property, known as noncontextual hidden-variable (NCHV) theories. Local hidden-variable theories, such as those discarded by Bell’s theorem, are a particular type of NCHV theories, so in this sense, the KS theorem is more general than Bell’s theorem.

3.3 The simplest proof of the KS theorem

The proof of the KS theorem with 18 projection operators in $\mathcal{H}_4$ [32, 33, 34] is given in table 1.

<table>
<thead>
<tr>
<th>1000</th>
<th>1111</th>
<th>1111</th>
<th>1000</th>
<th>1001</th>
<th>1001</th>
<th>1111</th>
<th>1111</th>
<th>1001</th>
</tr>
</thead>
<tbody>
<tr>
<td>0100</td>
<td>1111</td>
<td>1111</td>
<td>0010</td>
<td>0100</td>
<td>1111</td>
<td>1100</td>
<td>0101</td>
<td>0110</td>
</tr>
<tr>
<td>0011</td>
<td>1100</td>
<td>1010</td>
<td>0101</td>
<td>0010</td>
<td>1111</td>
<td>0011</td>
<td>1010</td>
<td>1111</td>
</tr>
<tr>
<td>0011</td>
<td>0011</td>
<td>0101</td>
<td>0101</td>
<td>1001</td>
<td>0110</td>
<td>1111</td>
<td>1111</td>
<td>1111</td>
</tr>
</tbody>
</table>

Table 1. The 18-vector roof of the KS theorem in $\mathcal{H}_4$.

Table 1 contains 18 vectors combined in 9 columns. Each vector appears twice. Each vector represents the projection operator onto the corresponding normalized vector. For instance, 0011 represents the projector onto the vector $\frac{1}{\sqrt{2}}(0, 0, 1, -1)$. Each column contains 4 mutually orthogonal vectors, so that the corresponding projectors sum the identity in $\mathcal{H}_4$. Therefore, in a NCHV theory, each column must have assigned the answer “yes” to one and only one vector. But it is easily
seen that such an assignment is impossible since each vector appears twice, so that the total number of "yes" answers must be an even number.

If we view $\mathcal{H}_4$ as a product of two tensor factors, $\mathcal{H}_2 \otimes \mathcal{H}_2$, corresponding to two qubits. Then, by realizing that the 18 vectors in table 1 are eigenvectors of some products of the usual representation of the Pauli matrices $\sigma_z$ and $\sigma_x$ for the spin state of spin-1/2 particles, we can rewritten table 1 as table 2.

<table>
<thead>
<tr>
<th>Table 2. The 18-vector proof of the KS theorem in $\mathcal{H}_2 \otimes \mathcal{H}_2$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$zz$   $xx$   $xx$   $zz$   $zzxx$   $zzxx$   $zxxz$   $zxxx$   $zzxz$   $zzxx$   $zzxx$   $zzxx$</td>
</tr>
<tr>
<td>$zx$   $zx$   $xx$   $zz$   $zxx$   $zx$   $xzx$   $xzx$   $zzz$   $zxz$   $zxx$   $zxx$   $zxxz$</td>
</tr>
<tr>
<td>$zx$   $zx$   $xx$   $zz$   $zxx$   $zx$   $xzx$   $xzx$   $zzz$   $zxz$   $zxx$   $zxx$   $zxxz$</td>
</tr>
<tr>
<td>$zx$   $zx$   $xx$   $zz$   $zxx$   $zx$   $xzx$   $xzx$   $zzz$   $zxz$   $zxx$   $zxx$   $zxxz$</td>
</tr>
</tbody>
</table>

The notation in table 2 is the following: $zx$ represents the yes-no question "are the spin component of first particle positive in the $z$ direction and the spin component of second particle negative in the $x$ direction?", and $zxxz$ denotes the yes-no question "are the products $zx := \sigma_1 z \otimes \sigma_2 x$ and $xz := \sigma_1 x \otimes \sigma_2 z$ negative and positive respectively?", etc. The first is an example of a factorizable yes-no question, since it can be answered after separate tests on the first and second particles. The latter is an example of an entangled yes-no question, since it cannot be answered after separate tests on both particles. Therefore, in table 2 there are two types of yes-no questions and, consequently, three types of maximal tests: those involving factorizable yes-no questions only, such as those in columns 1 to 4; those involving both factorizable and entangled yes-no questions, such as those in columns 4 to 8; and those involving entangled yes-no questions only, such as the one in the ninth column. Taking into account this hierarchy of experiments, the relevant elements of the proof of the KS theorem in $\mathcal{H}_2 \otimes \mathcal{H}_2$ can be illustrated as in fig. 1.

Beyond its graphical beauty, the interest of fig. 1 arises from the fact that it encapsulates the hierarchy of tests that is behind a gedanken experiment which challenged the old idea that the KS theorem could not be tested in a laboratory [38]. Such a gedanken experiment was refined in [39] and finally performed [40]. The connection between the 18-vector proof and the experiment proposed in [38] is explained in [34].

References

Fig. 1. Graphical representation of the 18-vector proof of the KS theorem. Each segment represents a yes-no question. Segments meeting in the same point are mutually compatible yes-no questions, and therefore each point represents a maximal test. Each of the 4 points delimiting a square in the centre is a test containing only factorizable yes-no questions. Each of the two points above and each of the two points below, is a test containing both factorizable and entangled yes-no questions. The point in the centre is a test containing only entangled yes-no questions.

10. A paper containing the calculations in this section has been submitted elsewhere.