arXiv:1201.0775v2 [quant-ph] 11 Jan 2012

On the Origins of Quantum Correlations

Joy Christian[1]
Wolfson College, University of Oxford,
Oxford, OX2 6UD, United Kingdom

Abstract: It is well known that quantum correlations are not only more disciplined (and hence stronger) compared to classical correlations, but they are more disciplined in a mathematically very precise sense. This raises an important physical question: What is responsible for making quantum correlations so much more disciplined? Here we explain the observed discipline of quantum correlations by identifying the symmetries of our physical space with those of a parallelized 7-sphere. We substantiate this identification by proving that any quantum correlation can be understood as a classical, local-realistic correlation among a set of points of a parallelized 7-sphere.

[1] jjc@alum.bu.edu

Chapter 1

On the Origins of Quantum Correlations

> In a good mystery story the most obvious clues often lead to the wrong suspects.
>
> Einstein and Infeld

1.1 Introduction

In 1927, when quantum theory was still in its infancy and John Bell was yet to be born, Albert Einstein—one of the founding fathers of the theory—was attending the now famous 5th Solvay Conference in Brussels. He was profoundly disturbed by what the new theory had to say about the nature of physical reality. Among other things, his concerns stemmed from a deep appreciation of unity in nature. Beyond the cliché of "God does not play dice", he had recognized that quantum theory entailed a fundamental schism in nature. The aphorism "God does not split reality" perhaps better captures the true essence of his concerns [1]. What he sought was a unified picture of nature, devoid of any subjective boundary between the classical and the quantum. What he suspected was a deeper layer of reality, beyond the polarized picture offered by quantum theory.

By 1935—when John Bell was seven—these worries of Einstein had matured into a powerful logical argument against the new theory. Published in collaboration with Boris Podolsky and Nathan Rosen, this argument proves, *once and for all*, that quantum theory provides at best an incomplete description of the physical reality [2]. Since

the argument itself is logically impeccable, this conclusion is beyond dispute. Any argument, however, can only be as good as its premises, and that—as is well known—is where Bell entered the game in 1964 [3]. He attempted to show that not all of the premises of EPR are mutually compatible. Ironically, however, it is the argument of Bell that turns out to contain a faulty assumption, not that of EPR. What is more, this assumption appears in the very first equation of Bell's famous paper [3], and yet it had escaped notice until recently.

In a series of papers, written between 2007 and 2011, I tried to bring out Bell's error and constructed explicit counterexamples, not only to his original theorem, but also to several of its variants. This book is a collection of these papers, each of which can be read more or less independently, but their contents are interconnected, and reveal different aspects of the fundamental flaw in Bell's argument. The collection as a whole, however, is better viewed as addressing a very important physical question. Regardless of the validity of his theorem, what Bell discovered in 1964 is physically quite significant. He discovered that quantum correlations are far more disciplined than any classically possible correlation. What is more, quantum correlations are not only more disciplined, but are more disciplined in a mathematically very precise sense. This tells us something much more profound about the structure of the world we live in. And, at the same time, it raises a very important physical question:

What is it that makes quantum correlations more disciplined than classical correlations?

My goal in this book is to answer this question in mathematically and physically as precise a sense as possible. To this end, let me begin with an extended summary of my argument against Bell's theorem.

As noted above, the story began with Einstein, Podolsky, and Rosen [2]. The logic of their argument can be summarized as follows:

$$
\begin{aligned}
& \textbf{(1)}\ \text{QM} \Longrightarrow \text{Perfect Correlations} \\
+\ & \textbf{(2)}\ \text{Adherence to Local Causality} \\
+\ & \textbf{(3)}\ \text{Criterion of Objective Reality} \\
+\ & \textbf{(4)}\ \text{Notion of a Complete Theory} \\
\Longrightarrow\ & \textbf{(5)}\ \text{QM is an Incomplete Theory.}
\end{aligned}
$$

Given their premises, the conclusion of EPR follows impeccably. Among their premises (which are hardly unreasonable), the one that concerns us the most is their criterion of completeness:

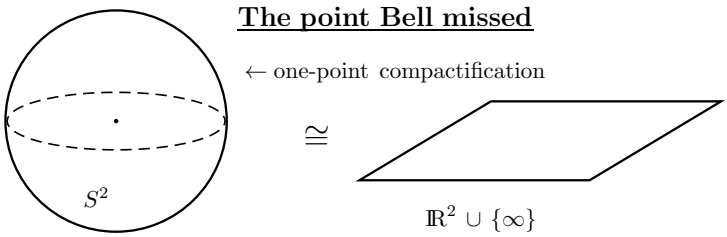

Figure 1.1: Although lines and planes contain the same number of points, it is impossible to put the points of a line or a plane in a one-to-one correspondence with all of the points of a 2-sphere.

every element of the physical reality must have a counterpart in the physical theory.

Bell attempted to prove that no theory satisfying this criterion can be locally causal. To this end, he took a complete theory to mean any theory whose predictions are dictated by functions of the form

$$\mathscr{A}(\mathbf{n},\, \lambda) : \mathbb{R}^3 \times \Lambda \longrightarrow S^0 \equiv \{-1,\, +1\}, \tag{1.1}$$

where $\mathbb{R}^3$ is the space of 3-vectors, Λ is a space of "complete" states, and $S^0 \equiv \{-1,\, +1\}$ is a unit 0-sphere. He then claimed (correctly, as it turns out) that no pair of functions of this form can reproduce the correlations for the singlet state predicted by quantum mechanics:

$$\langle\, \mathscr{A}(\mathbf{a},\, \lambda)\, \mathscr{B}(\mathbf{b},\, \lambda)\, \rangle \;\neq\; -\,\mathbf{a}\cdot\mathbf{b}\,. \tag{1.2}$$

At first sight, this appears to be a straightforward mathematical contradiction undermining the force of the EPR argument [3]. And for this reason functions of this form are routinely assumed in the Bell literature to provide complete specifications of the elements of physical reality, or complete accounting of all possible measurement results. As we shall see however, Bell's prescription is not only false, it is breathtakingly naïve and unphysical. It stems from an incorrect underpinning of both the EPR argument and the actual topological configurations involved in the relevant experiments [4]. In truth, for any two-level system the EPR criterion of completeness demands that the correct functions must necessarily be of the form

$$\pm\, 1 \;=\; \mathscr{A}(\mathbf{n},\, \lambda) : \mathbb{R}^3 \times \Lambda \longrightarrow S^3 \hookrightarrow \mathbb{R}^4, \tag{1.3}$$

with the *simply-connected* codomain S^3 of $\mathscr{A}(\mathbf{n},\,\lambda)$ replacing the *totally disconnected* codomain S^0 assumed by Bell. It is important to note here that this correction does not affect the actual measurement results. For a specific vector $\mathbf{n}$ and an initial state λ we still have

$$\mathscr{A}(\mathbf{n},\,\lambda)\;=\;+1\;\text{ or }\;-1 \tag{1.4}$$

as demanded by Bell, but now the topology of the codomain of the function $\mathscr{A}(\mathbf{n},\,\lambda)$ has changed from a 0-sphere to a 3-sphere, with the latter embedded in $\mathbb{R}^4$ in such a manner that the above constraint is satisfied. On the other hand, as is evident from Fig. (1.1) (and will be further clarified in the following pages), without this topological correction it is impossible to provide a complete account of all possible measurement results. Thus the selection of the codomain $S^3 \hookrightarrow \mathbb{R}^4$ in equation (1.3) is not a matter of choice but necessity. What is responsible for the EPR correlations is *the topology of the set of all possible measurement results* [4]. And for a two-level system this set happens to be an equatorial 2-sphere within a parallelized 3-sphere. But once the codomain of the functions $\mathscr{A}(\mathbf{n},\,\lambda)$ is so corrected, the proof of Bell's theorem (as given in Ref. [3]) simply falls apart. In fact, as we shall repeatedly see in the following pages, the strength of the correlation for *any* physical system is entirely determined by the topology of the codomain of the local functions $\mathscr{A}(\mathbf{n},\,\lambda)$. It has nothing whatsoever to do with entanglement or nonlocality.

1.2 Local Origins of the EPR-Bohm Correlations

Put differently, once the measurement results are represented by functions of the form (1.3), it is quite easy to reproduce the quantum correlations purely classically, in a manifestly local-realistic manner. For example, suppose Alice and Bob are equipped with the variables

$$\mathscr{A}(\mathbf{a},\,\lambda)\;=\;\{-\,a_j\,\boldsymbol{\beta}_j\,\}\,\{\,a_k\,\boldsymbol{\beta}_k(\lambda)\,\}\;=\;\begin{cases}+1 & \text{if}\quad \lambda\;=\;+1\\ -1 & \text{if}\quad \lambda\;=\;-1\end{cases} \tag{1.5}$$

and

$$\mathscr{B}(\mathbf{b},\,\lambda)\;=\;\{+\,b_k\,\boldsymbol{\beta}_k\,\}\,\{\,b_j\,\boldsymbol{\beta}_j(\lambda)\,\}\;=\;\begin{cases}-1 & \text{if}\quad \lambda\;=\;+1\\ +1 & \text{if}\quad \lambda\;=\;-1\,,\end{cases} \tag{1.6}$$

where the repeated indices are summed over x, y, and z; the fixed bivector basis $\{\boldsymbol{\beta}_x,\,\boldsymbol{\beta}_y,\,\boldsymbol{\beta}_z\}$ is defined by the algebra

$$\boldsymbol{\beta}_j\,\boldsymbol{\beta}_k\;=\;-\,\delta_{jk}\;-\;\epsilon_{jkl}\,\boldsymbol{\beta}_l\,; \tag{1.7}$$

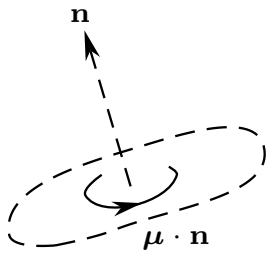

Figure 1.2: A unit bivector represents an equatorial point of a unit, parallelized 3-sphere. It is an abstraction of a directed plane segment, with only a magnitude and a sense of rotation—*i.e.*, clockwise $(-)$ or counterclockwise $(+)$. Neither the depicted oval shape of its plane, nor its axis of rotation $\mathbf{n}$, is an intrinsic part of the bivector $\boldsymbol{\mu}\cdot\mathbf{n}$.

and—together with $\boldsymbol{\beta}_j(\lambda) = \lambda\,\boldsymbol{\beta}_j$—the λ-dependent bivector basis $\{\boldsymbol{\beta}_x(\lambda),\, \boldsymbol{\beta}_y(\lambda),\, \boldsymbol{\beta}_z(\lambda)\}$ is defined by the algebra

$$\boldsymbol{\beta}_j\,\boldsymbol{\beta}_k \;=\; -\,\delta_{jk} \;-\; \lambda\,\epsilon_{jkl}\,\boldsymbol{\beta}_l\,, \tag{1.8}$$

where $\lambda = \pm 1$ is a fair coin representing two alternative orientations of the 3-sphere, δ_{jk} is the Kronecker delta, ϵ_{jkl} is the Levi-Civita symbol, and $\mathbf{a} = a_j\,\mathbf{e}_j$ and $\mathbf{b} = b_j\,\mathbf{e}_j$ are unit vectors [5]. Evidently, the variables $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$ belonging to S^3—in addition to being manifestly realistic—are strictly *local* variables. In fact, they are not even contextual [6]. Alice's measurement result—although it refers to a freely chosen direction $\mathbf{a}$—depends only on the initial state λ; and likewise, Bob's measurement result—although it refers to a freely chosen direction $\mathbf{b}$—depends only on the initial state λ.

In the subsequent chapters we shall mainly use the standard notations of Clifford algebra $Cl_{3,0}$. The bivector algebras (1.7) and (1.8) will then be seen as even subalgebras of $Cl_{3,0}$. The latter is a linear vector space, $\mathbb{R}^8$, spanned by the graded orthonormal basis

$$\{1,\;\mathbf{e}_x,\,\mathbf{e}_y,\,\mathbf{e}_z,\;\mathbf{e}_x\wedge\mathbf{e}_y,\,\mathbf{e}_y\wedge\mathbf{e}_z,\,\mathbf{e}_z\wedge\mathbf{e}_x,\;\mathbf{e}_x\wedge\mathbf{e}_y\wedge\mathbf{e}_z\}, \tag{1.9}$$

where "$\wedge$" is the outer product, and the trivector $I \equiv \mathbf{e}_x\wedge\mathbf{e}_y\wedge\mathbf{e}_z$ defines the fundamental volume form of the physical space. In terms of these notations we can rewrite the bivector $\{\,a_j\,\boldsymbol{\beta}_j(\lambda)\,\}$ as

$$\boldsymbol{\mu}\cdot\mathbf{a} \;\equiv\; \{\,a_j\,\boldsymbol{\beta}_j(\lambda)\,\} \;\equiv\; \lambda\,\{\,a_x\,\mathbf{e}_y\wedge\mathbf{e}_z \,+\, a_y\,\mathbf{e}_z\wedge\mathbf{e}_x \,+\, a_z\,\mathbf{e}_x\wedge\mathbf{e}_y\}, \tag{1.10}$$

with $\boldsymbol{\mu} = \lambda I$ now representing the hidden variable of the model. The variables $\mathscr{A}(\mathbf{a},\, \lambda)$ and $\mathscr{B}(\mathbf{b},\, \lambda)$ defined above then take the form

$$S^3 \ni \mathscr{A}(\mathbf{a},\, \boldsymbol{\mu}) \;=\; (-\,I\cdot\mathbf{a}\,)\,(+\,\boldsymbol{\mu}\cdot\mathbf{a}\,) \;=\; \begin{cases} +\,1 & \text{if} \quad \boldsymbol{\mu} \;=\; +\,I \\ -\,1 & \text{if} \quad \boldsymbol{\mu} \;=\; -\,I \end{cases} \tag{1.11}$$

and

$$S^3 \ni \mathscr{B}(\mathbf{b},\, \boldsymbol{\mu}) \;=\; (+\,I\cdot\mathbf{b}\,)\,(+\,\boldsymbol{\mu}\cdot\mathbf{b}\,) \;=\; \begin{cases} -\,1 & \text{if} \quad \boldsymbol{\mu} \;=\; +\,I \\ +\,1 & \text{if} \quad \boldsymbol{\mu} \;=\; -\,I, \end{cases} \tag{1.12}$$

with the trivector $\boldsymbol{\mu}$ being either $+\,I$ or $-\,I$ with equal probability. In what follows we shall view the fixed bivectors $(-\,I\cdot\mathbf{a}\,)$ and $(+\,I\cdot\mathbf{b}\,)$ as representing the measuring instruments for detecting the random bivectors $(+\,\boldsymbol{\mu}\cdot\mathbf{a}\,)$ and $(+\,\boldsymbol{\mu}\cdot\mathbf{b}\,)$, which represent the spins.

It is crucial to note that the variables $\mathscr{A}(\mathbf{a},\, \lambda)$ and $\mathscr{B}(\mathbf{b},\, \lambda)$ are generated with *different* bivectorial scales of dispersion (or different standard deviations) for each direction $\mathbf{a}$ and $\mathbf{b}$. Consequently, in statistical terms these variables are raw scores, as opposed to standard scores [7]. Recall that a standard score indicates how many standard deviations an observation or datum is above or below the mean. If x is a raw (or unnormalized) score and $\overline{\mathrm{x}}$ is its mean value, then the standard (or normalized) score, z(x), is defined by

$$\mathrm{z}(\mathrm{x}) \;=\; \frac{\mathrm{x} \,-\, \overline{\mathrm{x}}}{\sigma(\mathrm{x})}\,, \tag{1.13}$$

where $\sigma(\mathrm{x})$ is the standard deviation of x. A standard score thus represents the distance between a raw score and the population mean in the units of standard deviation, and allows us to make comparisons of raw scores that may have come from very different sources. In other words, the mean value of the standard score itself is always zero, with standard deviation unity. In terms of these concepts the bivariate correlation between raw scores x and y is defined as

$$\mathcal{E}(\mathrm{x},\, \mathrm{y}) \;=\; \frac{\displaystyle\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{i\,=\,1}^{n}\,(\mathrm{x}^i \,-\, \overline{\mathrm{x}})\,(\mathrm{y}^i \,-\, \overline{\mathrm{y}})\right]}{\sigma(\mathrm{x})\;\sigma(\mathrm{y})} \tag{1.14}$$

$$\;=\; \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{i\,=\,1}^{n}\,\mathrm{z}(\mathrm{x}^i)\;\mathrm{z}(\mathrm{y}^i)\right]. \tag{1.15}$$

It is vital to appreciate that covariance by itself—*i.e.*, the numerator of equation (1.14) by itself—does not provide the correct measure of

association between the raw scores, not the least because it depends on different units and scales (or different scales of dispersion) that may have been used (advertently or inadvertently) in the measurements of such scores [7]. Therefore, to arrive at the correct measure of association between the raw scores one must either use equation (1.14), with the product of standard deviations in the denominator, or use covariance of the standardized variables, as in Eq. (1.15).

These basic statistical concepts are crucial for understanding the EPR correlations. As defined above, the random variables $\mathscr{A}(\mathbf{a}, \lambda)$ and $\mathscr{B}(\mathbf{b}, \lambda)$ are products of two factors—one random and another non-random. Within $\mathscr{A}(\mathbf{a}, \lambda)$ the factor $\{ a_k \, \boldsymbol{\beta}_k(\lambda) \}$ is a random factor—a function of the hidden variable λ, whereas $\{ - a_j \, \boldsymbol{\beta}_j \}$ is a non-random factor, independent of the hidden variable λ. Thus, as a random variable each number $\mathscr{A}(\mathbf{a}, \lambda)$ and $\mathscr{B}(\mathbf{b}, \lambda)$ is generated with a *different* standard deviation—*i.e.*, a *different* size of typical error. More specifically, $\mathscr{A}(\mathbf{a}, \lambda)$ is generated with the standard deviation $\{ - a_j \, \boldsymbol{\beta}_j \}$, whereas $\mathscr{B}(\mathbf{b}, \lambda)$ is generated with a different standard deviation, namely $\{ + b_k \, \boldsymbol{\beta}_k \}$. These two deviations can be calculated easily. Since errors in linear relations propagate linearly, the standard deviation of $\mathscr{A}(\mathbf{a}, \lambda)$ is equal to $\{ - a_j \, \boldsymbol{\beta}_j \}$ times the standard deviation of $\{ a_k \, \boldsymbol{\beta}_k(\lambda) \}$ (which we write as $\sigma(A)$), whereas the standard deviation of $\mathscr{B}(\mathbf{b}, \lambda)$ is equal to $\{ + b_k \, \boldsymbol{\beta}_k \}$ times the standard deviation of $\{ b_j \, \boldsymbol{\beta}_j(\lambda) \}$ (which we write as $\sigma(B)$):

$$\begin{aligned} \sigma(\mathscr{A}) &= \{ - a_j \, \boldsymbol{\beta}_j \} \, \sigma(A) \\ \text{and} \quad \sigma(\mathscr{B}) &= \{ + b_k \, \boldsymbol{\beta}_k \} \, \sigma(B). \end{aligned} \tag{1.16}$$

But since all the bivectors we have been considering are normalized to unity, and since the mean of $\{ a_k \, \boldsymbol{\beta}_k(\lambda) \}$ vanishes on the account of λ being a fair coin, its standard deviation is easy to calculate, and it turns out to be equal to unity:

$$\begin{aligned} \sigma(A) &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left\| A(\mathbf{a}, \lambda^i) - \overline{A(\mathbf{a}, \lambda^i)} \right\|^2} \\ &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} \| \{ a_k \, \boldsymbol{\beta}_k(\lambda^i) \} - 0 \|^2} = 1, \end{aligned} \tag{1.17}$$

where the last equality follows from the fact that $\{ a_k \, \boldsymbol{\beta}_k(\lambda^i) \}$ are normalized to unity. Similarly, we find that $\sigma(B)$ is also equal to 1. As a result, the standard deviation of $\mathscr{A}(\mathbf{a}, \lambda)$ works out to be equal to $\{ - a_j \, \boldsymbol{\beta}_j \}$, and the standard deviation of $\mathscr{B}(\mathbf{b}, \lambda)$ works out to be equal to $\{ + b_k \, \boldsymbol{\beta}_k \}$. Putting these two results together, we arrive

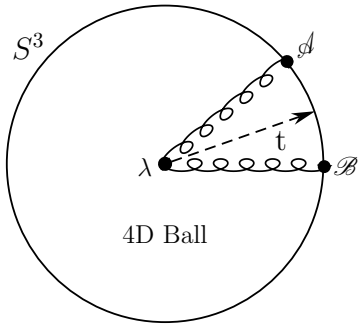

Figure 1.3: An initial EPR state λ originated at time $\mathrm{t} = 0$ evolves into measurement results $\mathscr{A}$ and $\mathscr{B}$ at a later time, occurring at two spacelike separated locations on a parallelized 3-sphere, S^3, which can be thought of as a boundary of a 4-dimensional ball of radius t. In what follows we shall assume that t has been normalized to unity.

at the following standardized scores corresponding to the raw scores:

$$\begin{aligned} A(\mathbf{a},\,\lambda) &= \frac{\mathscr{A}(\mathbf{a},\,\lambda)\;-\;\overline{\mathscr{A}(\mathbf{a},\,\lambda)}}{\sigma(\mathscr{A})} \\ &= \frac{\mathscr{A}(\mathbf{a},\,\lambda)\;-\;0}{\{-\,a_j\,\boldsymbol{\beta}_j\,\}} \;=\; \{\,a_k\,\boldsymbol{\beta}_k(\lambda)\,\} \end{aligned} \tag{1.18}$$

and

$$\begin{aligned} B(\mathbf{b},\,\lambda) &= \frac{\mathscr{B}(\mathbf{b},\,\lambda)\;-\;\overline{\mathscr{B}(\mathbf{b},\,\lambda)}}{\sigma(\mathscr{B})} \\ &= \frac{\mathscr{B}(\mathbf{b},\,\lambda)\;-\;0}{\{+\,b_k\,\boldsymbol{\beta}_k\,\}} \;=\; \{\,b_j\,\boldsymbol{\beta}_j(\lambda)\,\}, \end{aligned} \tag{1.19}$$

where we have used the identities such as $\{+\,a_k\,\boldsymbol{\beta}_k\,\}\{-\,a_k\,\boldsymbol{\beta}_k\,\} = +1$. Not surprisingly, just like the raw scores $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$, these standard scores are also strictly *local* variables: $A(\mathbf{a},\,\lambda)$ depends only on the freely chosen local direction $\mathbf{a}$ and the common cause λ, and likewise $B(\mathbf{b},\,\lambda)$ depends only on the freely chosen local direction $\mathbf{b}$ and the common cause λ. Moreover, despite appearances, $A(\mathbf{a},\,\lambda)$ and $B(\mathbf{b},\,\lambda)$ are simply binary numbers, $\pm\,1$, albeit occurring within

the compact topology of the 3-sphere rather than the real line:

$$S^3 \supset S^2 \ni A(\mathbf{a},\, \lambda) \;=\; \{\, a_k\, \boldsymbol{\beta}_k(\lambda)\,\} \;=\; \pm 1 \text{ about } \mathbf{a} \in \mathbb{R}^3, \tag{1.20}$$

$$S^3 \supset S^2 \ni B(\mathbf{b},\, \lambda) \;=\; \{\, b_j\, \boldsymbol{\beta}_j(\lambda)\,\} \;=\; \pm 1 \text{ about } \mathbf{b} \in \mathbb{R}^3. \tag{1.21}$$

In fact, since the space of all bivectors $\{\, a_k\, \boldsymbol{\beta}_k(\lambda)\,\}$ is isomorphic to the equatorial 2-sphere contained within the 3-sphere [4], each standard score $A(\mathbf{a},\, \lambda)$ of Alice is uniquely identified with a definite point of this 2-sphere, and likewise for the standard scores of Bob.

In the following chapters we shall tacitly assume that this procedure of standardizing from the raw scores to standard scores has been performed for all measurement results, taken either as points of a 3-sphere, or more generally as points of a 7-sphere.

Now, since we have assumed that initially there was 50/50 chance between the right-handed and left-handed orientations of the 3-sphere (*i.e.*, equal chance between the initial states $\lambda = +1$ and $\lambda = -1$), the expectation values of the local outcomes vanish trivially. On the other hand, as discussed above, to determine the *correct* correlation between the joint observations of Alice and Bob we must calculate covariance between the corresponding standard scores $A(\mathbf{a},\, \lambda)$ and $B(\mathbf{b},\, \lambda)$, not the raw scores themselves. The correlation between the raw scores $\mathscr{A}(\mathbf{a},\, \lambda)$ and $\mathscr{B}(\mathbf{b},\, \lambda)$ then works out to be

$$\mathcal{E}(\mathbf{a},\, \mathbf{b}) \;=\; \lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^{n} A(\mathbf{a},\, \lambda^i)\, B(\mathbf{b},\, \lambda^i) \right] \tag{1.22}$$

$$= \lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^{n} \{\, a_j\, \boldsymbol{\beta}_j(\lambda^i)\,\}\, \{\, b_k\, \boldsymbol{\beta}_k(\lambda^i)\,\} \right] \tag{1.23}$$

$$= -\, a_j\, b_j \;-\; \lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^{n} \{\, \lambda^i\, \epsilon_{jkl}\; a_j\, b_k\, \boldsymbol{\beta}_l\,\} \right] \tag{1.24}$$

$$= -\, a_j\, b_j \;+\; 0 \;=\; -\, \mathbf{a} \cdot \mathbf{b}\,, \tag{1.25}$$

where we have used the algebra defined in (1.8). Consequently, when the raw scores $\mathscr{A} = \pm 1$ and $\mathscr{B} = \pm 1$ are compared in practice by coincidence counts [8][9], the normalized expectation value of their product will inevitably yield

$$\mathcal{E}(\mathbf{a},\, \mathbf{b}) \;=\; \frac{\Big[C_{++}(\mathbf{a},\, \mathbf{b}) \;+\; C_{--}(\mathbf{a},\, \mathbf{b}) \;-\; C_{+-}(\mathbf{a},\, \mathbf{b}) \;-\; C_{-+}(\mathbf{a},\, \mathbf{b}) \Big]}{\Big[C_{++}(\mathbf{a},\, \mathbf{b}) \;+\; C_{--}(\mathbf{a},\, \mathbf{b}) \;+\; C_{+-}(\mathbf{a},\, \mathbf{b}) \;+\; C_{-+}(\mathbf{a},\, \mathbf{b}) \Big]}$$

$$= -\, \mathbf{a} \cdot \mathbf{b}\,, \tag{1.26}$$

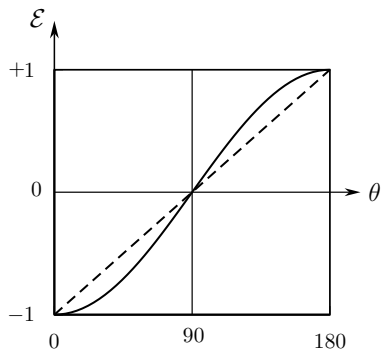

Figure 1.4: Local-realistic correlations can be stronger within S^3.

where $C_{+-}(\mathbf{a},\, \mathbf{b})$ *etc.* represent the number of joint occurrences of detections $+1$ along $\mathbf{a}$ and -1 along $\mathbf{b}$ *etc.*

The above equation simply describes covariance of the numbers $\mathscr{A} = \pm 1$ and $\mathscr{B} = \pm 1$ yielding the "impossible" strong correlation:

$$\mathcal{E}(\mathbf{a},\, \mathbf{b}) \;=\; \lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^{n} \mathscr{A}(\mathbf{a},\, \lambda^i)\, \mathscr{B}(\mathbf{b},\, \lambda^i) \right] \;=\; -\,\mathbf{a} \cdot \mathbf{b}\,. \qquad (1.27)$$

How is this "impossible" result possible? After all, Bell seems to have proved mathematically [3] that correlation between such numbers can never exceed the linear limit. The answer is that, apart from the statistical differences we discussed above, there are also topological differences between the above expression and what Bell considered in his theorem. In particular, in the above equation the numbers $\mathscr{A} = \pm 1$ and $\mathscr{B} = \pm 1$ are points of a parallelized 3-sphere,

$$\pm\, 1 \;=\; \mathscr{A}(\mathbf{n},\, \lambda) : \mathbb{R}^3 \times \Lambda \longrightarrow S^3 \hookrightarrow \mathbb{R}^4, \qquad (1.28)$$

and since a parallelized 3-sphere is a compact, simply-connected topological space, it should not be surprising that correlation among its points is stronger-than-linear. By contrast in Bell's argument the measurement functions are implicitly assumed to be of the form

$$\pm\, 1 \;=\; \mathscr{A}(\mathbf{n},\, \lambda) : \mathbb{R}^3 \times \Lambda \longrightarrow \mathcal{I} \subseteq \mathbb{R}\,, \qquad (1.29)$$

and since the topology of the real line is far less disciplined than that of a 3-sphere, it is impossible to generate strong correlation among its points. In other words, the strength of the correlation is

entirely dependent on how disciplined the codomain of the functions $\mathscr{A}(\mathbf{n},\,\lambda)$ is. Now it turns out that the parallelized spheres S^3 and S^7 have maximally disciplined topology in this respect [4]. And since the 3-sphere can be parallelized by unit quaternions, the covariance between its equatorial points represented by pure quaternions (or unit bivectors) is precisely the EPR correlation

$$\mathcal{E}(\mathbf{a},\,\mathbf{b}) \;=\; \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{i\,=\,1}^{n} A(\mathbf{a},\,\lambda^i)\;B(\mathbf{b},\,\lambda^i)\right] \;=\; -\,\mathbf{a}\cdot\mathbf{b}\,, \qquad (1.30)$$

where the bivectors $A(\mathbf{a},\,\lambda) = \{\,a_k\;\boldsymbol{\beta}_k(\lambda)\}$ and $B(\mathbf{b},\,\lambda) = \{\,b_k\;\boldsymbol{\beta}_k(\lambda)\}$ are the standardized variables in the statistical terms discussed above. Then, for arbitrary four directions $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$, and $\mathbf{b}'$, the corresponding CHSH string of expectation values immediately gives

$$\begin{aligned}&|\mathcal{E}(\mathbf{a},\,\mathbf{b}) \,+\, \mathcal{E}(\mathbf{a},\,\mathbf{b}') \,+\, \mathcal{E}(\mathbf{a}',\,\mathbf{b}) \,-\, \mathcal{E}(\mathbf{a}',\,\mathbf{b}')| \\ &\qquad\qquad \leq\; 2\,\sqrt{1-(\mathbf{a}\times\mathbf{a}')\cdot(\mathbf{b}'\times\mathbf{b})} \;\leq\; 2\sqrt{2}\,, \qquad (1.31)\end{aligned}$$

which squarely contradicts Bell's theorem and exactly reproduces the quantum mechanical prediction. So much for Bell's theorem.

1.3 Local Origins of ALL Quantum Correlations

What the above example shows is that the correlation among the raw scores observed by Alice and Bob are entirely determined by the topology of the codomain of the local functions $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$. In particular, correlations among the points of a unit parallelized 3-sphere are stronger than those among the points of the real line. Moreover, as we shall see in the following chapters, once the topology of the codomain is correctly specified, not only the EPR correlations, but also the correlations predicted by the rotationally non-invariant quantum states—such as the GHZ states and Hardy state—can be exactly reproduced in a purely local-realistic manner. Thus, contrary to the widespread belief, the correlations exhibited by such states are not some irreducible quantum effects, but purely local-realistic, topological effects. In the cases of EPR and Hardy states the correct topology of the codomain is a parallelized 3-sphere, and consequently the correlations exhibited by these states are classical correlations among the points of a parallelized 3-sphere. In the case of GHZ states on the other hand the correct topology of the codomain is a parallelized 7-sphere, and consequently the correlations exhibited by these states are classical correlations among the points of a parallelized 7-sphere. More generally, all quantum mechanical correlations can

be understood as purely classical, local-realistic correlations among the points of a parallelized 7-sphere. Needless to say, this vindicates Einstein's suspicion that a quantum state merely describes a statistical ensemble of physical systems, not an individual physical system.

What is more, as we shall see in Chapter 7 in greater detail, there are profound mathematical and conceptual reasons why the topology of the 7-sphere plays such a significant role in the manifestation of *all* quantum correlations. As is well known, quantum correlations are more disciplined (or stronger) than classical correlations in a mathematically precise sense. It turns out that it is the discipline of parallelization in the manifold of all possible measurement results that is responsible for the strength of quantum correlations. In fact, both the existence as well as the strength of all quantum correlations are dictated by the parallelizability of the spheres S^0, S^1, S^3, and S^7, with the 7-sphere being homeomorphic to the most general possible division algebra. And it is the property of division that is responsible for maintaining the strict local causality in the world we live in. These considerations lead us to the following remarkable theorem:

<u>Theorema Egregium:</u>

Every quantum mechanical correlation can be understood as a classical, local-realistic correlation among a set of points of a parallelized 7-sphere, represented by maps of the form

$$\pm\, 1 \,=\, \mathscr{A}(\mathbf{n},\, \lambda) : \mathbb{R}^3 \times \Lambda \longrightarrow S^7 \hookrightarrow \mathbb{R}^8. \tag{1.32}$$

The corresponding physical picture is the same as in Figure 1.3, but with S^3 replaced by S^7, the 4D ball replaced by the 8D ball, and the number of statistically significant measurement events, $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$, *etc.*, generalized to an arbitrary number. It is important to note, however, that despite appearances neither S^3 nor S^7 is a round sphere. The Riemann curvature of both S^3 and S^7 is zero, because they are both parallelized spheres [4].

S^7, however, has a much richer topological structure than S^3. As noted above, it happens to be homeomorphic to the space of unit octonions, which are well known to form the most general division algebra possible. In the language of fiber bundles one can thus view a 7-sphere as a 4-sphere worth of 3-spheres. Each of its fiber is then a 3-sphere, and each one of these 3-spheres is a 2-sphere worth of circles. Thus the four parallelizable spheres—S^0, S^1, S^3, and S^7—can all be viewed as nested within a 7-sphere. The EPR-Bohm correlations can then be understood as correlations among the equatorial points of one of the fibers of this 7-sphere, as we saw above. Alternatively, a 7-sphere can be thought of as a 6-sphere worth of circles. Thus the above theorem can be framed entirely in terms of circles, each one of

which described by a classical, octonionic spinor with a well-defined sense of rotation (i.e., whether it describes a clockwise rotation about a point within the 7-sphere or a counterclockwise rotation). This sense of rotation in turn defines a definite handedness (or orientation) about every point of the 7-sphere. If we designate this handedness by a random number $\lambda = \pm 1$, then local measurement results for any physical scenario can be represented by raw scores of the form

$$S^7 \ni \mathscr{A}(\mathbf{a},\, \boldsymbol{\mu}) \;=\; (-\,J \cdot \mathbf{N}(\mathbf{a})\,)\,(+\,\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a})\,) \;=\; \begin{cases} +\,1 & \text{if} \;\; \boldsymbol{\mu} \;=\; +\,J \\ -\,1 & \text{if} \;\; \boldsymbol{\mu} \;=\; -\,J, \end{cases} \tag{1.33}$$

where $\mathbf{a} \in \mathbb{R}^3$ and $\mathbf{N}(\mathbf{a}) \in \mathbb{R}^7$ are unit vectors, and $\boldsymbol{\mu} = \lambda\, J$ is the hidden variable analogous to $\boldsymbol{\mu} = \lambda\, I$ with $I = \mathbf{e}_x\mathbf{e}_y\mathbf{e}_z$ replaced by

$$J = \;\mathbf{e}_1\mathbf{e}_2\mathbf{e}_4 + \mathbf{e}_2\mathbf{e}_3\mathbf{e}_5 + \mathbf{e}_3\mathbf{e}_4\mathbf{e}_6 + \mathbf{e}_4\mathbf{e}_5\mathbf{e}_7 + \mathbf{e}_5\mathbf{e}_6\mathbf{e}_1 + \mathbf{e}_6\mathbf{e}_7\mathbf{e}_2 + \mathbf{e}_7\mathbf{e}_1\mathbf{e}_3. \tag{1.34}$$

The standard scores corresponding to these raw scores are then given by $\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a})$, which geometrically represent the equatorial points of a parallelized 7-sphere, just as $\boldsymbol{\mu} \cdot \mathbf{a}$ represented the equatorial points of a parallelized 3-sphere. In Chapters 6 and 7 we shall see explicit examples of the functions $\mathbf{N}(\mathbf{a}) \in \mathbb{R}^{\mathbf{7}}$ corresponding to measurement directions $\mathbf{a} \in \mathbb{R}^3$. The correlations between the raw scores are then calculated as covariance between the standard scores $\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a})$. Note also that, just as in the EPR case, both the raw scores (1.33) and the standard scores $\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a})$ are manifestly non-contextual.

Equipped with these variables, we are now in a position to prove the above theorem [4]. To this end, recall that no matter which model of physics we are concerned with—the quantum mechanical model, the hidden variable model, or any other—for theoretical purposes all we need to understand are the expectation values of the observables measured in various states of the physical systems [10]. Accordingly, consider an arbitrary quantum state $|\Psi\rangle \in \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space of arbitrary dimensions, which may or may not be finite. We impose no restrictions on either $|\Psi\rangle$ or $\mathcal{H}$, apart from their usual quantum mechanical meanings. In particular, the state $|\Psi\rangle$ can be as entangled as one may like, and the space $\mathcal{H}$ can be as large or small as one may like. Next consider a self-adjoint operator $\widehat{\mathcal{O}}(\mathbf{a},\, \mathbf{b},\, \mathbf{c},\, \mathbf{d},\, \ldots)$ on this Hilbert space, parameterized by an arbitrary number of local contexts $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, *etc.* The quantum mechanical expectation value of this observable in the state $|\Psi\rangle$ is then given by:

$$\mathcal{E}_{Q.M.}(\mathbf{a},\, \mathbf{b},\, \mathbf{c},\, \mathbf{d},\, \ldots) \;=\; \langle\Psi|\, \widehat{\mathcal{O}}(\mathbf{a},\, \mathbf{b},\, \mathbf{c},\, \mathbf{d},\, \ldots)\, |\Psi\rangle\,. \tag{1.35}$$

More generally, if the system happens to be in a mixed state, then

$$\mathcal{E}_{Q.M.}(\mathbf{a},\, \mathbf{b},\, \mathbf{c},\, \mathbf{d},\, \ldots) \;=\; \mathrm{Tr}\left\{W\, \widehat{\mathcal{O}}(\mathbf{a},\, \mathbf{b},\, \mathbf{c},\, \mathbf{d},\, \ldots)\right\}, \tag{1.36}$$

where W is a statistical operator of unit trace representing the state.

Our goal now is to show that this expectation value can always be reproduced as a local-realistic expectation value of a set of binary points of a parallelized 7-sphere. To this end, let

$$\pm\,1 \;=\; \mathscr{A}_{\mathbf{a}}(\lambda) : \mathbb{R}^3 \times \Lambda \longrightarrow S^7, \quad \pm 1 \;=\; \mathscr{B}_{\mathbf{b}}(\lambda) : \mathbb{R}^3 \times \Lambda \longrightarrow S^7, \;\; etc. \tag{1.37}$$

be the raw scores of the form (1.33). Using prescriptions analogous to (1.18) the corresponding standard scores then work out to be

$$\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a}) : \mathbb{R}^3 \times \Lambda \longrightarrow S^6, \quad \boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{b}) : \mathbb{R}^3 \times \Lambda \longrightarrow S^6, \;\; etc. \tag{1.38}$$

Here $\mathbf{N}(\mathbf{a})$, $\mathbf{N}(\mathbf{b})$, *etc.* may not necessarily be the same function for all $\mathbf{n} \in \mathbb{R}^3$. They may be different functions for different directions.

This idealized prescription of raw scores and standard scores can, and should, be further generalized. So far we have presumed that randomness in these scores originates entirely from the initial state λ representing the orientation of the 7-sphere. In other words, we have presumed that the local interactions of the measuring devices $(-\,J \cdot \mathbf{N}(\mathbf{a})\,)$ with the physical variables $\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a})$ do not introduce additional randomness in the scores $\mathscr{A}_{\mathbf{a}}(\lambda)$. Any realistic interaction between $(-\,J \cdot \mathbf{N}(\mathbf{a})\,)$ and $\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a})$, however, *would* introduce such a randomness of purely local origin. We can model it by an additional random variable $\nu = \pm 1$ with probability $0 \leq p(\,\nu \,|\, \mathbf{a},\, \lambda\,) \leq 1$, so that the bivectors $(-\,J \cdot \mathbf{N}(\mathbf{a})\,)$ representing the measuring devices may now also take the random form $(-\,\nu\, J \cdot \mathbf{N}(\mathbf{a})\,)$. The average of the corresponding raw scores $\mathscr{A}_{\mathbf{a}}(\lambda) = \pm 1$ would then satisfy

$$-\,1 \;\leq\; \overline{\mathscr{A}_{\mathbf{a}}(\lambda)} \;\leq\; +1\,, \tag{1.39}$$

$$\begin{aligned} \text{with} \quad \overline{\mathscr{A}_{\mathbf{a}}(\lambda)} \;&=\; \sum_{\nu} p(\,\nu \,|\, \mathbf{a},\, \lambda\,)\, \mathscr{A}_{\mathbf{a}}(\nu,\, \lambda) \\ &=\; \left[\sum_{\nu} p(\,\nu \,|\, \mathbf{a},\, \lambda\,)\, (-\,\nu\, J \cdot \mathbf{N}(\mathbf{a})\,) \right] (+\,\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a})\,) \\ &=\; \overline{(-\,J \cdot \mathbf{N}(\mathbf{a})\,)}\, (+\,\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a})\,)\,. \end{aligned} \tag{1.40}$$

Not surprisingly, this does not affect the corresponding standard scores (1.38) worked out earlier. But if, in addition, we assume that the common cause $\lambda = \pm 1$ itself is also distributed non-uniformly between its values $+1$ or -1, then the standard scores modify to

$$\boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a}) \longrightarrow \kappa(\lambda)\, \boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{a}), \quad \boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{b}) \longrightarrow \kappa(\lambda)\, \boldsymbol{\mu} \cdot \mathbf{N}(\mathbf{b}) \;\; etc., \tag{1.41}$$

where
$$\kappa(\lambda) \;=\; \frac{1-\lambda\,\overline{\lambda}}{\sqrt{1-\left(\overline{\lambda}\right)^2}}\,, \tag{1.42}$$

with $\overline{\lambda}$ being the average over the probability distribution of λ. The correlation among the raw scores $\mathscr{A}_{\mathbf{a}}(\lambda)=\pm 1$, $\mathscr{B}_{\mathbf{b}}(\lambda)=\pm 1$, $\mathscr{C}_{\mathbf{c}}(\lambda)=\pm 1$, *etc.* can now be easily calculated as covariance among the standard scores $A_{\mathbf{a}}(\lambda)=\boldsymbol{\mu}\cdot\mathbf{N}(\mathbf{a})$, $B_{\mathbf{b}}(\lambda)=\boldsymbol{\mu}\cdot\mathbf{N}(\mathbf{b})$, *etc.* as

$$\mathcal{E}_{L.R.}(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\,\ldots) \;=\; \int_{\Lambda} A_{\mathbf{a}}(\lambda)\,B_{\mathbf{b}}(\lambda)\,C_{\mathbf{c}}(\lambda)\,\ldots\;\;\rho(\lambda)\,d\lambda\,, \tag{1.43}$$

where the overall probability distribution $\rho(\lambda)=\kappa^m(\lambda)$ is allowed to be both a non-uniform and continuous function of λ, with m being the total number of local contexts $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, $\ldots$ in the experiment.

To evaluate these correlations, note that the standard scores $A_{\mathbf{a}}(\lambda)=\boldsymbol{\mu}\cdot\mathbf{N}(\mathbf{a})$, $B_{\mathbf{b}}(\lambda)=\boldsymbol{\mu}\cdot\mathbf{N}(\mathbf{b})$, $C_{\mathbf{c}}(\lambda)=\boldsymbol{\mu}\cdot\mathbf{N}(\mathbf{c})$, *etc.* are in fact bivectors representing the equatorial points of the 7-sphere, which remains as closed under multiplication as the 3-sphere. As a result, the product of the standard scores can be written as

$$A_{\mathbf{a}}(\lambda)\,B_{\mathbf{b}}(\lambda)\,C_{\mathbf{c}}(\lambda)\cdots = f(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\ldots)+P_{\widehat{\mathbf{N}}}(\lambda)\;\;g\,(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\ldots)\,, \tag{1.44}$$

where the RHS is an octonionic spinor representing a non-equatorial point of the 7-sphere, the vector $\widehat{\mathbf{N}}(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\,\ldots)\in\mathbb{R}^7$ is a function of all 3D vectors, $P_{\widehat{\mathbf{N}}}(\lambda)\equiv\boldsymbol{\mu}\cdot\widehat{\mathbf{N}}(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\ldots)$ is a unit bivector representing an equatorial point of S^7 that is different from the ones represented by the bivectors $A_{\mathbf{a}}(\lambda)$, $B_{\mathbf{b}}(\lambda)$, $C_{\mathbf{c}}(\lambda)$, *etc.*, and the scalar functions $f(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\ldots)$ and $g(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\ldots)$ satisfy $f^2+g^2=1$ with f identified as the quantum mechanical expectation value (1.36):

$$f(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\,\ldots) \;=\; \mathrm{Tr}\left\{W\,\widehat{\mathcal{O}}(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\,\ldots)\right\}. \tag{1.45}$$

Conversely, any arbitrary point of the 7-sphere (or joint beable)

$$(\,A_{\mathbf{a}}\,B_{\mathbf{b}}\,C_{\mathbf{c}}\,D_{\mathbf{d}}\ldots)(\lambda) \;=\; f(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\ldots)+P_{\widehat{\mathbf{N}}}(\lambda)\;\;g\,(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\ldots) \tag{1.46}$$

corresponding to the quantum mechanical operator $\widehat{\mathcal{O}}(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\ldots)$ can always be factorized into any number of local parts as

$$S^7 \;\ni\; (\,A_{\mathbf{a}}\,B_{\mathbf{b}}\,C_{\mathbf{c}}\,D_{\mathbf{d}}\;\ldots)(\lambda) \;=\; A_{\mathbf{a}}(\lambda)\,B_{\mathbf{b}}(\lambda)\,C_{\mathbf{c}}(\lambda)\,D_{\mathbf{d}}(\lambda)\;\ldots\,, \tag{1.47}$$

since, as we have already noted, the 7-sphere remains closed under multiplication of any number of its points. Using the identity (1.44),

the local realistic expectation value (1.43) can now be rewritten as

$$\mathcal{E}_{L.R.}(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\,\ldots) \;=\; f(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\,\ldots)\int_{\Lambda}\rho(\lambda)\;d\lambda$$
$$+\;g\,(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\,\ldots)\int_{\Lambda}P_{\widehat{\mathbf{N}}(\mathbf{n})}(\lambda)\;\;\rho(\lambda)\;d\lambda\,. \quad (1.48)$$

Note, however, that the vector $\widehat{\mathbf{N}}(\mathbf{n}) \in \mathbb{R}^7$ appearing in the second term here corresponds to a 3D direction $\mathbf{n} \in \mathbb{R}^3$ that is necessarily exclusive to all the other measurement directions $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}, \ldots$ As a result, the bivector $P_{\widehat{\mathbf{N}}(\mathbf{n})}(\lambda) = \boldsymbol{\mu}\cdot\widehat{\mathbf{N}}(\mathbf{n})$ necessarily corresponds to a null measurement result, reducing the second integral in (1.48) to zero [4][11]. If we next assume that the probability distribution, although not necessarily uniform, remains normalized to unity,

$$\int_{\Lambda}\rho(\lambda)\;d\lambda \;=\; 1\,, \quad (1.49)$$

then the above expectation value reduces to

$$\mathcal{E}_{L.R.}(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\,\ldots) \;=\; f(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d},\,\ldots)\,. \quad (1.50)$$

This finally proves our main theorem: Every quantum mechanical correlation can be understood as a classical, local-realistic correlation among a set of points of a parallelized 7-sphere. Q.E.D.

1.4 The Raison D'être of Quantum Correlations

The above result demonstrates that the discipline of parallelization in the manifold of all possible measurement results is responsible for the existence and strength of *all* quantum correlations. More precisely, it identifies quantum correlations as evidence that the physical space we live in respects the symmetries and topologies of a parallelized 7-sphere. As we shall see in greater detail in Chapter 7, there are profound mathematical and conceptual reasons why the topology of the 7-sphere plays such a significant role in the manifestation of quantum correlations. Essentially it is because 7-sphere happens to be homeomorphic to the most general possible division algebra. And it is the property of division that turns out to be responsible for maintaining strict local causality in the world we live in.

To understand this chain of reasoning better, recall that, just as a parallelized 3-sphere is a 2-sphere worth of 1-spheres but with a twist in the manifold S^3 ($\neq S^2 \times S^1$), a parallelized 7-sphere is a 4-sphere worth of 3-spheres but with a twist in the manifold S^7 ($\neq S^4 \times S^3$).

More precisely, just as S^3 is a nontrivial fiber bundle over S^2 with Clifford parallels S^1 as its linked fibers, S^7 is also a nontrivial fiber bundle, but over S^4, and with entire 3-dimensional spheres S^3 as its linked fibers. Now it is the twist in the bundle S^3 that forces one to forgo the commutativity of complex numbers (corresponding to the circles S^1) in favor of the non-commutativity of quaternions. In other words, a 3-sphere is not parallelizable by the commuting complex numbers but only by the non-commuting quaternions. In a similar vein, the twist in the bundle $S^7 \neq S^4 \times S^3$ forces one to forgo the associativity of quaternions (corresponding to the fibers S^3) in favor of the non-associativity of octonions. In other words, a 7-sphere is not parallelizable by the associative quaternions but only by the non-associative octonions. And the reason why it can be parallelized at all is because its tangent bundle happens to be trivial:

$$\mathrm{T}S^7 = \bigcup_{p \in S^7} \{p\} \times T_p S^7 \equiv S^7 \times \mathbb{R}^7. \tag{1.51}$$

Once parallelized by a set of unit octonions, both the 7-sphere and each of its 3-spherical fibers remain closed under multiplication. This, in turn, means that the factorizability or locality condition of Bell is automatically satisfied within a parallelized 7-sphere. The lack of associativity of octonions, however, entails that, unlike the unit 3-sphere (which is homeomorphic to the Lie group SU(2)), a 7-sphere is not a group manifold, but forms only a quasi-group. As a result, the torsion within the 7-sphere continuously varies from one point to another of the manifold [4]. It is this variability of the parallelizing torsion within S^7 that is ultimately responsible for the diversity and non-linearity of the quantum correlations we observe in nature:

$$\text{Parallelizing Torsion } \mathcal{T}^{\gamma}_{\alpha\,\beta} \neq 0 \quad \Longleftrightarrow \quad \text{Quantum Correlations.}$$

The upper bound on all possible quantum correlations is thus set by the maximum of possible torsion within the 7-sphere:

$$\text{Maximum of Torsion } \mathcal{T}^{\gamma}_{\alpha\,\beta} \neq 0 \quad \Longrightarrow \quad \text{The Upper Bound } 2\sqrt{2}.$$

These last two results will be proved rigorously in Chapter 7.

1.5 Local Causality and the Division Algebras

In the last few sections we saw the crucial role played by the 3- and 7-dimensional spheres in understanding the existence of quantum

correlations. What is so special about 3 and 7 dimensions? Why is the vector cross product definable only in 3 and 7 dimensions and no other? Why are $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$ the only possible normed division algebras? Why are only the 3- and 7-dimensional spheres nontrivially parallelizable out of infinitely many possible spheres? Why is it possible to derive all quantum mechanical correlations as local-realistic correlations among the points of only the 7-sphere?

The answers to all of these questions are intimately connected to the notion of factorizability introduced by Bell within the context of his theorem. Mathematicians have long been asking: When is a product of two squares itself a square: $x^2\,y^2 = z^2$? If the number z is factorizable, then it can be written as a product of two other numbers, $z = x\,y$, and then the above equality is seen to hold for the numbers x, y, and z. For ordinary numbers this is easy to check. The number 8 can be factorized into a product of 2 and 4, and we then have $64 = 8^2 = (2 \times 4)^2 = 2^2 \times 4^2 = 64$. But what about sums of squares? A more profound equality holds, in fact, for a sum of two squares times a sum of two squares as a third sum of two squares:

$$(x_1^2 + x_2^2)\,(y_1^2 + y_2^2) \;=\; (x_1 y_1 - x_2 y_2)^2 + (x_1 y_2 + x_2 y_1)^2 \;=\; z_1^2 + z_2^2. \tag{1.52}$$

There is also an identity like this one for the sums of four squares. It was first discovered by Euler, and then rediscovered and popularized by Hamilton in the 19th century through his work on quaternions. It is also known that Graves and Cayley independently discovered a similar identity for the sums of eight squares. This naturally leads to the question of whether the product of two sums of squares of n different numbers can be a sum of n different squares? In other words, does the following equality hold in general for any n?

$$(x_1^2 + x_2^2 + \cdots + x_n^2)\,(y_1^2 + y_2^2 + \cdots + y_n^2) \;=\; z_1^2 + z_2^2 + \cdots + z_n^2. \tag{1.53}$$

It turns out that this equality holds only for $n = 1$, 2, 4, and 8. This was proved by Hurwitz in 1898 [13]. It reveals a deep and surprising fact about the world we live in. Much of what we see around us, from elementary particles to distant galaxies, is an inevitable consequence of this simple mathematical fact. The world is the way it is because the above equality holds only for $n = 1$, 2, 4, and 8. For example, the above identity is equivalent to the existence of a division algebra of dimension n over the field $\mathbb{R}$ of real numbers. Indeed, if we define vectors $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$, and $\mathbf{z} = \mathbf{x} * \mathbf{y}$ in $\mathbb{R}^n$ such that z_i's are functions of x_j's and y_k's determined by (1.53), then

$$||\mathbf{x}||\,||\mathbf{y}|| \;=\; ||\mathbf{x} * \mathbf{y}||\,. \tag{1.54}$$

Thus the division algebras $\mathbb{R}$ (real), $\mathbb{C}$ (complex), $\mathbb{H}$ (quaternion), and $\mathbb{O}$ (octonion) we use in much of our science are intimately related

to the dimensions $n = 1$, 2, 4, and 8. Moreover, from the equation of a unit sphere,

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 1\,, \tag{1.55}$$

it is easy to see that the four parallelizable spheres S^0, S^1, S^3, and S^7 correspond to $n = 1$, 2, 4, and 8, which are the dimensions of the respective embedding spaces of these four spheres. What is not so easy to see, however, is the fact that there is a deep connection between Hurwitz's theorem and the quantum correlations [4]. As we saw in the previous sections, all quantum correlations are inevitable consequences of the parallelizability of the 7-sphere, which in turn is a consequence of Hurwitz's theorem. So the innocent looking algebraic equality (1.53) has far reaching consequences, not only for the entire edifice of mathematics, but also for that of quantum physics.

1.6 Concluding Remarks

The results obtained in the preceding sections (and in the chapters that are to follow [4]) go far deeper and well beyond the boundaries of Bell's theorem. In any physical experiment what is observed are "clicks" of the event detectors corresponding to yes/no answers to our questions. As in the EPR experiment [9][12], when we compare the answers recorded by various observers in quantum experiments conducted at mutually remote locations, we find that their answers are correlated in a mathematically and statistically very disciplined manner. The natural question then clearly is: why are these answers correlated in such a disciplined manner when there appears to be no predetermined common cause dictating the correlations. Bell and his followers claimed that the observed correlations are the evidence of a radical non-locality in nature. The stronger adherents of this view often claim that "nature is non-local." Here we have rejected this view. Instead, we have demonstrated how a perfectly natural local explanation of the observed correlations is possible. In our view these correlations are the evidence, not of non-locality, but the fact that the physical space we live in respects the symmetries and topologies of a parallelized 7-sphere [4]. The 3-dimensional space we normally presume as our reality is then simply one of the many fibers of this 7-sphere. Our observations are still confined to various 3-dimensional subspaces of this 7-sphere, but the correlations among the results of our experiments are revealing that the observed measurement events are actually occurring within a 7-dimensional manifold. The radical non-locality of Bell is thus traded off for the extra dimensions going beyond our immediate experiences in the macroscopic world.

The methodology that has led us to this conclusion is similar to

that used by Einstein to arrive at his local field theory of gravity. Just as the demand of locality in the face of Newton's non-local theory of gravity led Einstein to general relativity, the demand of locality in the face of quantum correlations has led us to a parallelized 7-sphere.

Acknowledgments

I wish to thank David Coutts, Fred Diether, Azhar Iqbal, Edwin Eugene Klingman, Rick Lockyer, Ray B. Munroe, and Tom H. Ray for discussions. I also wish to thank the Foundational Questions Institute (FQXi) for supporting this work through a Mini-Grant.

Appendix: The Meaning of a Geometric Product

The concept of a geometric product was first introduced by Hermann Grassmann to characterize what he called "extensive magnitudes." Nowadays Hestenes refers to "extensive magnitudes" as "directed numbers" [5]. To understand the meaning of the geometric product between two such directed numbers, let us first look at the "inner product" between two vectors, say $\mathbf{a}$ and $\mathbf{b}$:

$$\mathbf{a}\cdot\mathbf{b} = \frac{1}{2}(\mathbf{a}\,\mathbf{b} + \mathbf{b}\,\mathbf{a}) = \cos\theta_{\mathbf{ab}} = \mathbf{b}\cdot\mathbf{a}\,, \tag{1.56}$$

where $\theta_{\mathbf{ab}}$ is the angle between $\mathbf{a}$ and $\mathbf{b}$. Clearly, this product is a grade-*lowering* operation. It takes two grade-1 numbers, or vectors, and gives back a grade-0 number, or a scalar.

Next, let us look at the "outer product" between $\mathbf{a}$ and $\mathbf{b}$, as defined by Grassmann:

$$\mathbf{a}\wedge\mathbf{b} = \frac{1}{2}(\mathbf{a}\,\mathbf{b} - \mathbf{b}\,\mathbf{a}) = I\cdot(\mathbf{a}\times\mathbf{b}) = -\,\mathbf{b}\wedge\mathbf{a}\,, \tag{1.57}$$

where I (in the modern parlance) is a trivector. Unlike the previous product this product is a grade-*raising* operation. It takes two grade-1 numbers, or vectors, and gives back a new directed number of grade 2; *i.e.*, a bivector. Although an abstract entity, the bivector $\mathbf{a}\wedge\mathbf{b}$ may be visualized as an oriented plane segment, hovering orthogonal to the vector $\mathbf{a}\times\mathbf{b}$.

Using the products (1.56) and (1.57), the geometric product between the two directed numbers $\mathbf{a}$ and $\mathbf{b}$ can now be expressed as

$$\boxed{\mathbf{a}\,\mathbf{b} = \mathbf{a}\cdot\mathbf{b} + \mathbf{a}\wedge\mathbf{b}}\,. \tag{1.58}$$

This product also takes two grade-1 numbers, or vectors, but gives back an entity that is neither a scalar nor a bivector in general, but

rather a quaternion (or a spinor), made out of the grade-lowering operation $\mathbf{a} \cdot \mathbf{b}$ and the grade-raising operation $\mathbf{a} \wedge \mathbf{b}$. To appreciate this, let us express the two components of the geometric product (1.58) more explicitly as

$$\begin{aligned} \mathbf{a}\,\mathbf{b} &= \mathbf{a} \cdot \mathbf{b} + I \cdot (\mathbf{a} \times \mathbf{b}) \\ &= \cos\theta_{\mathbf{ab}} + (\, I \cdot \mathbf{c}\,) \sin\theta_{\mathbf{ab}} \\ &= \exp\{(\, I \cdot \mathbf{c}\,)\, \theta_{\mathbf{ab}}\}\,, \end{aligned} \tag{1.59}$$

where $\mathbf{c} = \mathbf{a} \times \mathbf{b}/|\mathbf{a} \times \mathbf{b}|$. It is now clear that the product $\mathbf{a}\,\mathbf{b}$ is a quaternion (or a spinor) that represents a rotation by an angle $2\theta_{\mathbf{ab}}$ about the $\mathbf{c}$-axis. The geometric product, $\mathbf{a}\,\mathbf{b}$, thus takes two grade-1 numbers, or vectors $\mathbf{a}$ and $\mathbf{b}$, and gives back *a pure act of rotation.*

References

[1] A. Einstein, Dialectica **2**, 320 (1948).

[2] A. Einstein, B. Podolsky, N. Rosen, Phys. Rev. **47**, 777 (1935).

[3] J. S. Bell, Physics **1**, 195 (1964).

[4] J. Christian, *What Really Sets the Upper Bound on Quantum Correlations?*, arXiv:1101.1958

[5] D. Hestenes and G. Sobczyk, *Clifford Algebra to Geometric Calculus* (Reidel, Dordrecht, 1984); T. G. Vold, Am. J. Phys, **61**, 491 (1993); D. Hestenes, *New Foundations for Classical Mechanics*, Second Edition (Kluwer, Dordrecht, 1999); D. Hestenes, Am. J. Phys, **71**, 104 (2003); C. Doran and A. Lasenby, *Geometric Algebra for Physicists* (Cambridge University Press, 2003).

[6] A. Shimony, Brit. J. Phil. Sci. **35**, 25 (1984).

[7] J. L. Rodgers and W. A. Nicewander, The American Statistician **42**, 59 (1988).

[8] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **49**, 91 (1982); See also A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).

[9] G. Weihs *et al.*, Phys. Rev. Lett. **81**, 5039 (1998).

[10] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955).

[11] J. Christian, *Disproofs of Bell, GHZ, and Hardy Type Theorems and the Illusion of Entanglement*, arXiv:0904.4259

[12] J. Christian, *Disproof of Bell's Theorem*, arXiv:1103.1879.

[13] A. Hurwitz, Nachr. Ges. Wiss. Göttingen 309 (1898).