Bell's Theorem Begs the Question

Joy Christian

Einstein Centre for Local-Realistic Physics, Oxford OX2 6LB, United Kingdom; jjc@bu.edu

Abstract: I demonstrate that Bell's theorem is based on circular reasoning and thus a fundamentally flawed argument. It unjustifiably assumes the additivity of expectation values for dispersion-free states of contextual hidden variable theories for non-commuting observables involved in Bell-test experiments, which is tautologous to assuming the bounds of ± 2 on the Bell-CHSH sum of expectation values. Its premises thus assume in a different guise the bounds of ± 2 it sets out to prove. Once this oversight is ameliorated from Bell's argument by identifying the impediment that leads to it and local realism is implemented correctly, the bounds on the Bell-CHSH sum of expectation values work out to be $\pm 2\sqrt{2}$ instead of ± 2, thereby mitigating the conclusion of Bell's theorem. Consequently, what is ruled out by any of the Bell-test experiments is not local realism but the linear additivity of expectation values, which does not hold for non-commuting observables in any hidden variable theories to begin with.

Keywords: Bell's theorem; local realism; Bell-CHSH inequalities; quantum correlations; Bell-test experiments

---

1. Introduction

Bell's theorem [1] is an impossibility argument (or "proof") that claims that no locally causal and realistic hidden variable theory envisaged by Einstein [2] that could "complete" quantum theory can reproduce all of the predictions of quantum theory. But some such claims of impossibility in physics are known to harbor unjustified assumptions. In this paper, I show that Bell's theorem against locally causal hidden variable theories is no exception. It is no different, in this respect, from von Neumann's theorem against all hidden variable theories [3], or the Coleman-Mandula theorem overlooking the possibilities of supersymmetry [4]. The implicit and unjustified assumptions underlying the latter two theorems seemed so innocuous to many that they escaped notice for decades. By contrast, Bell's theorem has faced skepticism and challenges by many from its very inception (cf. footnote 1 in [5]), including by me [5–16], because it depends on a number of questionable implicit and explicit physical assumptions that are not difficult to recognize [10,16]. In what follows, I bring out one such assumption and demonstrate that Bell's theorem is based on a circular argument [9]. It unjustifiably assumes the additivity of expectation values for dispersion-free states of hidden variable theories for non-commuting observables involved in the Bell-test experiments [17], which is tautologous to assuming the bounds of ± 2 on the Bell-CHSH sum of expectation values. Its premises thus assume in a different guise what it sets out to prove. Once this oversight is ameliorated from Bell's argument, the local-realistic bounds on the Bell-CHSH sum of expectation values work out to be $\pm 2\sqrt{2}$ instead of ± 2, thereby mitigating the conclusion of Bell's theorem. As a result, what is ruled out by the Bell-test experiments is not local realism but the additivity of expectation values, which does not hold for non-commuting observables in dispersion-free states of hidden variable theories to begin with.

1.1. Heuristics for Completing Quantum Mechanics

The goal of any hidden variable theory [3,18,19] is to reproduce the statistical predictions encoded in the quantum states $|\psi\rangle \in \mathscr{H}$ of physical systems using hypothetical dispersion-free states $|\psi, \lambda\rangle := \{|\psi\rangle, \lambda\} \in \mathscr{H} \otimes \mathscr{L}$ that have no inherent statistical character, where the Hilbert space $\mathscr{H}$ is extended by the space $\mathscr{L}$ of hidden variables λ, which are hypothesized to "complete" the states of the physical

systems as envisaged by Einstein [2]. If the values of $\lambda \in \mathscr{L}$ can be specified in advance, then the results of any measurements on a given physical system are uniquely determined.

To appreciate this, recall that expectation value of the square of any self-adjoint operator $\Omega \in \mathscr{H}$ in a normalized quantum mechanical state $|\psi\rangle$ and the square of the expectation value of Ω will not be equal to each other in general:

$$\langle\psi|\,\Omega^2\,|\psi\rangle \neq \langle\psi\,|\,\Omega\,|\,\psi\rangle^2. \tag{1}$$

This gives rise to inherent statistical uncertainty in the value of Ω, indicating that the state $|\psi\rangle$ is not dispersion-free:

$$\Delta\Omega = \sqrt{\langle\psi|\{\,\Omega - \langle\psi|\,\Omega\,|\psi\rangle\mathbb{1}\}^2\,|\psi\rangle} \neq 0. \tag{2}$$

By contrast, in a normalized dispersion-free state $|\psi,\lambda)$ of hidden variable theories formalized by von Neumann [3], the expectation value of Ω, *by hypothesis*, is equal to one of its eigenvalues $\omega(\lambda)$, determined by the hidden variables λ,

$$(\,\psi,\lambda\,|\,\Omega\,|\,\psi,\lambda\,) = \omega(\lambda) \iff \Omega\,|\,\psi,\lambda) = \omega(\lambda)\,|\,\psi,\lambda), \tag{3}$$

so that a measurement of Ω in the state $|\,\psi,\lambda\,)$ would yield the result $\omega(\lambda)$ with certainty. How this can be accomplished in a dynamical theory of measurement process remains an open question [18]. But accepting the hypothesis (3) implies

$$(\psi,\lambda\,|\,\Omega^2\,|\,\psi,\lambda) = (\psi,\lambda\,|\,\Omega\,|\,\psi,\lambda)^2. \tag{4}$$

Consequently, unlike in a quantum sate $|\psi\rangle$, in a dispersion-free state $|\psi,\lambda)$ observables Ω have no inherent uncertainty:

$$\Delta\Omega = \sqrt{(\,\psi,\lambda\,|\,\{\,\Omega - (\,\psi,\lambda\,|\,\Omega\,|\,\psi,\lambda\,)\,\mathbb{1}\}^2\,|\,\psi,\lambda)} = 0. \tag{5}$$

The expectation value of Ω in the quantum state $|\psi\rangle$ can then be recovered by integrating over the hidden variables λ:

$$\langle\,\psi\,|\,\Omega\,|\,\psi\,\rangle = \int_{\mathscr{L}} (\,\psi,\lambda\,|\,\Omega\,|\,\psi,\lambda\,)\,p(\lambda)\,d\lambda = \int_{\mathscr{L}} \omega(\lambda)\,p(\lambda)\,d\lambda, \tag{6}$$

where $p(\lambda)$ denotes the normalized probability distribution over the space $\mathscr{L}$ of thus hypothesized hidden variables. The quantum mechanical dispersion (2) in the measured value of the observer Ω can thus be interpreted as due to the distribution $p(\lambda)$ in the values of the hidden variables λ over the statistical ensemble of the physical systems measured.

As it stands, this prescription amounts to assignment of unique eigenvalues $\omega(\lambda)$ to *all* observables Ω *simultaneously*, regardless of whether they are actually measured. In other words, according to (6) every physical quantity of a given system represented by Ω would possess a unique preexisting value, irrespective of any measurements being performed. The prescription (6) thus mathematically encodes Einstein's conception of realism. In [2], Einstein explained his point of view in terms of the position and momentum of a free particle as follows: "The (free) particle really has a definite position and a definite momentum, even if they cannot both be ascertained by measurement in the same individual case [as in quantum mechanics]. According to this point of view, the ψ-function represents an incomplete description of the real state of affairs." The acceptance of this point of view — Einstein continues — "would lead to an attempt to obtain a complete description of the real state of affairs *as well as the incomplete one* (my emphasis), and to discover physical laws for such a description." Accordingly, the left-hand side of the first equality in (6) provides the incomplete description of the system and its right-hand side provides the complete one, with all possible statements one can make about the system encoded in the expectation values of

the observables being measured in the state of the system [3]. If Ω_1 and Ω_2 are two non-commuting observables, then the uncertainty relation between them can also be interpreted as

$$\Delta\Omega_1 \, \Delta\Omega_2 \; \geqslant \; \frac{1}{2} \, |\langle\,\psi\,|\,[\Omega_1,\Omega_2]\,|\,\psi\,\rangle| = \frac{1}{2} \left| \int_{\mathscr{L}} (\,\psi,\lambda\,|\,[\Omega_1,\Omega_2]\,|\,\psi,\lambda\,) \; p(\lambda)\,d\lambda \right| = \frac{1}{2} \left| \int_{\mathscr{L}} \omega_{1,2}(\lambda) \; p(\lambda)\,d\lambda \right| \tag{7}$$

in terms of the probability distribution $p(\lambda)$ in the values of the hidden variables λ, where the first inequality is the one established by Robertson [20], and $i\,\omega_{1,2}(\lambda)$ is a purely imaginary eigenvalue of the skew-Hermitian operator $[\Omega_1,\Omega_2]$. Similarly, using the kinematical equivalence seen in (6), the dynamical equivalence between the quantum mechanical description and Einstein's "complete" description can also be established, as demonstrated in the Appendix A below:

$$\left[\frac{d}{dt} \langle\,\psi\,|\,\Omega\,|\,\psi\,\rangle = \frac{1}{i\hbar} \langle\,\psi\,|\,[\Omega,H]\,|\,\psi\,\rangle + \langle\,\psi\,|\,\frac{\partial\Omega}{\partial t}\,|\,\psi\,\rangle \right] = \int_{\mathscr{L}} \left[\frac{d}{dt}\,\omega(\lambda) = \{\omega(\lambda), \mathscr{H}(\lambda)\} + \frac{\partial\,\omega(\lambda)}{\partial t} \right] p(\lambda)\,d\lambda\,. \tag{8}$$

If particular values of λ are precisely known with $p(\lambda) = 1$, then the right-hand side of (8) would reduce to the classical Hamiltonian equations of motion. Otherwise, Ehrenfest's equation in quantum mechanics on the left-hand side of (8) can be understood as an ensemble average of classical dynamics over the probability distribution $p(\lambda)$ of λ. Thus, each probabilistic statement about the system can be traced back to the incompleteness of our knowledge about the system.

In Section 2 of [18], Bell works out an instructive example to illustrate how the prescription (6) works for a system of two-dimensional Hilbert space. It fails, however, for Hilbert spaces of dimensions greater than two, because in higher dimensions degeneracies prevent simultaneous assignments of unique eigenvalues to all observables in dispersion-free states $|\,\psi,\lambda\,\rangle$ dictated by the ansatz (3), giving contradictory values for the same physical quantities. This was proved independently by Bell [18], Kochen and Specker [21], and Belinfante [22], as a corollary to Gleason's theorem [23,24].

These proofs – known as the Kochen-Specker theorem – do not exclude contextual hidden variable theories in which the complete state $|\,\psi,\lambda\rangle$ of a system assigns unique values to physical quantities only *relative* to experimental contexts [19,24]. If we denote the observables as $\Omega(c)$ with c being the environmental contexts of their measurements, then the non-contextual prescription (6) can be easily modified to accommodate contextual hidden variable theories as follows:

$$\langle\,\psi\,|\,\Omega(c)\,|\,\psi\,\rangle = \int_{\mathscr{L}} (\,\psi,\lambda\,|\,\Omega(c)\,|\,\psi,\lambda\,) \; p(\lambda)\,d\lambda = \int_{\mathscr{L}} \omega(c,\lambda) \; p(\lambda)\,d\lambda\,. \tag{9}$$

Each observable $\Omega(c)$ is still assigned a unique eigenvalue $\omega(c,\lambda)$, but now determined cooperatively by the complete state $|\,\psi,\lambda\,\rangle$ of the system and the state c of its environmental contexts. Consequently, even though some of its features are no longer intrinsic to the system, contextual hidden variable theories do not have the inherent statistical character of quantum mechanics, because outcome of an experiment is a cooperative effect just as it is in classical physics [24]. Therefore, such theories interpret quantum entanglement at the level of the complete state $|\,\psi,\lambda\,\rangle$ only epistemically.

For our purposes here, it is also important to recall that in the Hilbert space formulation of quantum mechanics [3] the correspondence between observables and Hermitian operators is one-to-one. Moreover, a sum $\widetilde{\Omega}(\tilde{c}) = \sum_{i=1}^{n}\Omega_i(c_i)$ of several observables such as $\Omega_1(c_1), \Omega_2(c_2), \Omega_3(c_3),\ldots,\Omega_n(c_n)$ is also an observable representing a physical quantity, and consequently the sum of the expectation values of $\Omega_i(c_i)$ is the expectation value of the summed operator $\widetilde{\Omega}(\tilde{c})$,

$$\sum_{i=1}^{n} \langle\,\psi\,|\,\Omega_i(c_i)\,|\,\psi\,\rangle = \langle\,\psi\,|\left[\sum_{i=1}^{n}\Omega_i(c_i)\right]|\,\psi\,\rangle, \tag{10}$$

regardless of whether the observables are simultaneously measurable or mutually commutative [18]. The question then is, since within any contextual hidden variable theory characterized by (9) all of the observables $\Omega_i(c_i)$ and their sum $\widetilde{\Omega}(\tilde{c})$ are assigned unique eigenvalues $\omega_i(c_i, \lambda)$ and $\widetilde{\omega}(\tilde{c}, \lambda)$, respectively, would these eigenvalues satisfy the equality

$$\sum_{i=1}^n \left[\int_{\mathscr{L}} \omega_i(c_i, \lambda) \, p(\lambda) \, d\lambda \right] \overset{?}{=} \int_{\mathscr{L}} \left[\sum_{i=1}^n \omega_i(c_i, \lambda) \right] p(\lambda) \, d\lambda \tag{11}$$

in dispersion-free states $|\psi, \lambda\rangle$ of physical systems in analogy with the linear quantum mechanical relation (10) above? The answer is: Not in general, because the eigenvalue $\widetilde{\omega}(\tilde{c}, \lambda)$ of the summed operator $\widetilde{\Omega}(\tilde{c})$ is not equal to the sum $\sum_{i=1}^n \omega_i(c_i, \lambda)$ of eigenvalues $\omega_i(c_i, \lambda)$ for given λ, unless the constituent observables $\Omega_i(c_i)$ are mutually commutative. In other words, $\widetilde{\omega}(\tilde{c}, \lambda) \neq \sum_{i=1}^n \omega_i(c_i, \lambda)$ in general, and therefore the correct counterpart of relation (10) is not (11) but

$$\sum_{i=1}^n \left[\int_{\mathscr{L}} \omega_i(c_i, \lambda) \, p(\lambda) \, d\lambda \right] = \int_{\mathscr{L}} \widetilde{\omega}(\tilde{c}, \lambda) \, p(\lambda) \, d\lambda \,. \tag{12}$$

As Bell points out in Section 3 of [18], the linear relation (10) is an unusual property of quantum mechanical states $|\psi\rangle$. There is no reason to demand it *individually* of the dispersion-free states $|\psi, \lambda\rangle$, whose function is to reproduce the measurable features of quantum systems only when averaged over, as in (9). There is no reason why the value of $\Omega(c)$ should not be determined by some *nonlinear* function $(\psi, \lambda \,|\, \Omega(c) \,|\, \psi, \lambda)$. I will come back to this issue in Section 1.4.

In [18], Bell explains this non-linearity using spin components of a spin-$\frac{1}{2}$ particle. Suppose we make a measurement of the component σ_x of the spin with a Stern-Gerlach magnet suitably oriented in $\mathbb{R}^3$. That would yield an eigenvalue s_x of σ_x as a result. However, if we wish to measure the component σ_y of the spin, then that would require a different orientation of the magnet in $\mathbb{R}^3$, and would give a different eigenvalue, s_y of σ_y, as a result. Moreover, a measurement of the sum of the x- and y-components of the spin, $\sigma_x + \sigma_y$, would again require a very different orientation of the magnet in $\mathbb{R}^3$. Therefore, the result obtained as an eigenvalue of the summed operators $\sigma_x + \sigma_y$ will not be the sum $s_x + s_y$ of an eigenvalue of the operator σ_x added linearly to an eigenvalue of the operator σ_y. Indeed, the eigenvalues of σ_x and σ_y are both ± 1, while the eigenvalues of $\sigma_x + \sigma_y$ are $\pm\sqrt{2}$, so a linear relation cannot hold. As Bell points out in [18], the additivity of expectation values $\langle \psi \,|\, \sigma_x \,|\, \psi \rangle + \langle \psi \,|\, \sigma_y \,|\, \psi \rangle = \langle \psi \,|\, \sigma_x + \sigma_y \,|\, \psi \rangle$ is a rather unusual property of the quantum states $|\psi\rangle$. The linearity of it is effectuated in quantum mechanics by promoting observable quantities to self-adjoint operators [25]. It does not hold for the dispersion-free states $|\psi, \lambda\rangle$ of hidden variable theories in general because the eigenvalues of non-commuting observables such as σ_x and σ_y do not add linearly, as we noted above. Consequently, the additivity relation (10) that holds for quantum states would not hold for the dispersion-free states.

1.2. Special Case of the Singlet State and EPR-Bohm Observables

Now, the proof of Bell's famous theorem [1] is based on Bohm's spin version of the EPR's thought experiment [26], which involves an entangled pair of spin-$\frac{1}{2}$ particles emerging from a source and moving freely in opposite directions, with particles 1 and 2 subject, respectively, to spin measurements along independently chosen unit directions **a** and **b** by Alice and Bob, who are stationed at a spacelike separated distance from each other (see Fig. 1). If initially the pair has vanishing total spin, then the quantum mechanical state of the system is described by the entangled singlet state

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \Big\{ |\mathbf{k}, +\rangle_1 \otimes |\mathbf{k}, -\rangle_2 - |\mathbf{k}, -\rangle_1 \otimes |\mathbf{k}, +\rangle_2 \Big\}, \tag{13}$$

where $\mathbf{k}$ is an arbitrary unit vector in $\mathbb{R}^3$ and

$$\boldsymbol{\sigma} \cdot \mathbf{k} \, |\mathbf{k}, \pm\rangle = \pm |\mathbf{k}, \pm\rangle \tag{14}$$

defines quantum mechanical eigenstates in which the two fermions have spins "up" or "down" in the units of $\hbar = 2$, with $\boldsymbol{\sigma}$ being the Pauli spin "vector" $(\sigma_x, \sigma_y, \sigma_z)$. Once the state (13) is prepared, the observable $\Omega(c)$ of interest is

$$\Omega(c) = \boldsymbol{\sigma}_1 \cdot \mathbf{a} \otimes \boldsymbol{\sigma}_2 \cdot \mathbf{b}, \tag{15}$$

whose possible eigenvalues are

$$\omega(c, \lambda) = \mathscr{A}\mathscr{B}(\mathbf{a}, \mathbf{b}, \lambda) = \pm 1, \tag{16}$$

where $\mathscr{A} = \pm 1$ and $\mathscr{B} = \pm 1$ are the results of spin measurements made jointly by Alice and Bob along their randomly chosen detector directions $\mathbf{a}$ and $\mathbf{b}$. In the singlet state (13), the joint observable (15) predicts sinusoidal correlations $\langle \Psi | \boldsymbol{\sigma}_1 \cdot \mathbf{a} \otimes \boldsymbol{\sigma}_2 \cdot \mathbf{b} | \Psi \rangle = -\mathbf{a} \cdot \mathbf{b}$ between the values of the spins observed about the freely chosen contexts $\mathbf{a}$ and $\mathbf{b}$ [6].
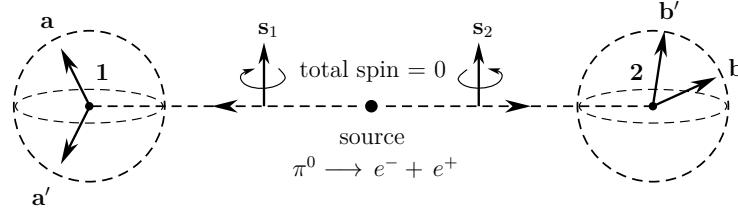
Figure 1. In an EPR-Bohm-type experiment, a spin-less fermion – such as a neutral pion – is assumed to decay from a source into an electron-positron pair, as depicted. Then, measurements of the spin components of each separated fermion are performed at space-like separated observation stations **1** and **2**, obtaining binary results $\mathscr{A} = \pm 1$ and $\mathscr{B} = \pm 1$ along directions $\mathbf{a}$ and $\mathbf{b}$. The conservation of spin momentum dictates that the total spin of the system remains zero during its free evolution. After Ref. [5].

For *locally* contextual hidden variable theories there is a further requirement that the results of local measurements must be describable by functions that respect local causality, as first envisaged by Einstein [2] and later formulated mathematically by Bell [1]. It can be satisfied by requiring that the eigenvalue $\omega(c, \lambda)$ of the observable $\Omega(c)$ in (15) representing the joint result $\mathscr{A}\mathscr{B}(\mathbf{a}, \mathbf{b}, \lambda) = \pm 1$ is factorizable as $\omega(c, \lambda) = \omega_1(c_1, \lambda) \, \omega_2(c_2, \lambda)$, or in Bell's notation as

$$\mathscr{A}\mathscr{B}(\mathbf{a}, \mathbf{b}, \lambda) = \mathscr{A}(\mathbf{a}, \lambda) \, \mathscr{B}(\mathbf{b}, \lambda), \tag{17}$$

with the factorized functions $\mathscr{A}(\mathbf{a}, \lambda) = \pm 1$ and $\mathscr{B}(\mathbf{b}, \lambda) = \pm 1$ satisfying the following condition of local causality:

> Apart from the hidden variables λ, the result $\mathscr{A} = \pm 1$ of Alice depends *only* on the measurement context $\mathbf{a}$, chosen freely by Alice, regardless of Bob's actions [1]. And, likewise, apart from the hidden variables λ, the result $\mathscr{B} = \pm 1$ of Bob depends *only* on the measurement context $\mathbf{b}$, chosen freely by Bob, regardless of Alice's actions. In particular, the function $\mathscr{A}(\mathbf{a}, \lambda)$ *does not* depend on $\mathbf{b}$ or $\mathscr{B}$ and the function $\mathscr{B}(\mathbf{b}, \lambda)$ *does not* depend on $\mathbf{a}$ or $\mathscr{A}$. Moreover, the hidden variables λ do not depend on either $\mathbf{a}$, $\mathbf{b}$, $\mathscr{A}$, or $\mathscr{B}$ [11].

The expectation value $\mathcal{E}(\mathbf{a}, \mathbf{b})$ of the joint results in the dispersion-free state $|\psi, \lambda\rangle$ should then satisfy the condition

$$\langle \Psi | \, \boldsymbol{\sigma}_1 \cdot \mathbf{a} \otimes \boldsymbol{\sigma}_2 \cdot \mathbf{b} \, | \Psi \rangle = \mathcal{E}(\mathbf{a}, \mathbf{b}) := \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}, \lambda) \, \mathscr{B}(\mathbf{b}, \lambda) \, p(\lambda) \, d\lambda, \tag{18}$$

where the hidden variables λ originate from a source located in the overlap of the backward light cones of Alice and Bob, and the normalized probability distribution $p(\lambda)$ is assumed to remain statistically independent of the contexts $\mathbf{a}$ and $\mathbf{b}$ so that $p(\lambda \,|\, \mathbf{a}, \mathbf{b}) = p(\lambda)$, which is a reasonable assumption. In fact, relaxing this assumption to allow $p(\lambda)$ to depend on $\mathbf{a}$ and $\mathbf{b}$ introduces a form of non-locality, as explained by Clauser and Horne in footnote 13 of [27]. Then, since $\mathscr{A}(\mathbf{a}, \lambda) = \pm 1$ and $\mathscr{B}(\mathbf{b}, \lambda) = \pm 1$, their product $\mathscr{A}(\mathbf{a}, \lambda)\,\mathscr{B}(\mathbf{b}, \lambda) = \pm 1$, setting the following bounds on $\mathcal{E}(\mathbf{a}, \mathbf{b})$:

$$-1 \leqslant \mathcal{E}(\mathbf{a}, \mathbf{b}) \leqslant +1. \tag{19}$$

These bounds are respected not only by local hidden variable theories but also by quantum mechanics and experiments.

1.3. Mathematical Core of Bell's Theorem

By contrast, at the heart of Bell's theorem is a derivation of the bounds of ± 2 on an *ad hoc* sum of the expectation values $\mathcal{E}(\mathbf{a}, \mathbf{b})$ of local results $\mathscr{A}(\mathbf{a}, \lambda)$ and $\mathscr{B}(\mathbf{b}, \lambda)$, recorded at remote observation stations by Alice and Bob, from four different sub-experiments involving measurements of non-commuting observables such as $\sigma_1 \cdot \mathbf{a}$ and $\sigma_1 \cdot \mathbf{a}'$ [1,17]:

$$\mathcal{E}(\mathbf{a}, \mathbf{b}) + \mathcal{E}(\mathbf{a}, \mathbf{b}') + \mathcal{E}(\mathbf{a}', \mathbf{b}) - \mathcal{E}(\mathbf{a}', \mathbf{b}')\,. \tag{20}$$

Alice can freely choose a detector direction $\mathbf{a}$ or $\mathbf{a}'$, and likewise Bob can freely choose a detector direction $\mathbf{b}$ or $\mathbf{b}'$, to detect, at a space-like distance from each other, the spins of fermions they receive from the common source. Then, from (19), we can immediately read off the upper and lower bounds on the combination (20) of expectation values:

$$-4 \leqslant \mathcal{E}(\mathbf{a}, \mathbf{b}) + \mathcal{E}(\mathbf{a}, \mathbf{b}') + \mathcal{E}(\mathbf{a}', \mathbf{b}) - \mathcal{E}(\mathbf{a}', \mathbf{b}') \leqslant +4\,. \tag{21}$$

1.3.1. Standard Derivation of the Bell-CHSH Inequalities (25)

The next step in Bell's derivation of the bounds ± 2 instead of ± 4 is the assumption of the additivity of expectation values:

$$\mathcal{E}(\mathbf{a}, \mathbf{b}) + \mathcal{E}(\mathbf{a}, \mathbf{b}') + \mathcal{E}(\mathbf{a}', \mathbf{b}) - \mathcal{E}(\mathbf{a}', \mathbf{b}')$$
$$= \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}, \lambda)\mathscr{B}(\mathbf{b}, \lambda)\, p(\lambda)\, d\lambda + \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}, \lambda)\mathscr{B}(\mathbf{b}', \lambda)\, p(\lambda)\, d\lambda + \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}', \lambda)\mathscr{B}(\mathbf{b}, \lambda)\, p(\lambda)\, d\lambda - \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}', \lambda)\mathscr{B}(\mathbf{b}', \lambda)\, p(\lambda)\, d\lambda$$
$$= \int_{\mathscr{L}} \big\{\, \mathscr{A}(\mathbf{a}, \lambda)\,\mathscr{B}(\mathbf{b}, \lambda) + \mathscr{A}(\mathbf{a}, \lambda)\,\mathscr{B}(\mathbf{b}', \lambda) + \mathscr{A}(\mathbf{a}', \lambda)\,\mathscr{B}(\mathbf{b}, \lambda) - \mathscr{A}(\mathbf{a}', \lambda)\,\mathscr{B}(\mathbf{b}', \lambda) \,\big\}\, p(\lambda)\, d\lambda\,. \tag{22}$$

We will have much to discuss about this seemingly harmless mathematical step employing the built-in linear additivity of integrals. I will demonstrate in Section 1.4 that, far from being harmless, it is, in fact, *an unjustified assumption that harbors a profound mistake of assuming the very thesis of the theorem to be proven*, just as von Neumann's theorem did [9,18,25,28,29]. It assumes, *without proof*, that linear additivity of integrals leading to (22) can be meaningfully applied, not only to the eigenvalues of commuting observables but also to the eigenvalues of non-commuting observables that cannot be measured simultaneously. But, as we saw in the paragraph following equation (12), this assumption is quite mistaken. However, if we overlook this mistake and accept equality (22), then the bounds of ± 2 on the Bell-CHSH combination (20) of expectation values are not difficult to work out by rewriting the integrand on its right-hand side as

$$\mathscr{A}(\mathbf{a}, \lambda)\,\big\{\, \mathscr{B}(\mathbf{b}, \lambda) + \mathscr{B}(\mathbf{b}', \lambda) \,\big\} + \mathscr{A}(\mathbf{a}', \lambda)\,\big\{\, \mathscr{B}(\mathbf{b}, \lambda) - \mathscr{B}(\mathbf{b}', \lambda) \,\big\}\,. \tag{23}$$

Since $\mathscr{B}(\mathbf{b}, \lambda) = \pm 1$, if $|\mathscr{B}(\mathbf{b}, \lambda) + \mathscr{B}(\mathbf{b}', \lambda)| = 2$, then $|\mathscr{B}(\mathbf{b}, \lambda) - \mathscr{B}(\mathbf{b}', \lambda)| = 0$, and vice versa. Consequently, since $\mathscr{A}(\mathbf{a}, \lambda) = \pm 1$, the integrand (23) is bounded by ± 2 and the absolute value of the last integral in (22) does not exceed 2:

$$-2 \leqslant \int_{\mathscr{L}} \left\{ \mathscr{A}(\mathbf{a}, \lambda) \mathscr{B}(\mathbf{b}, \lambda) + \mathscr{A}(\mathbf{a}, \lambda) \mathscr{B}(\mathbf{b}', \lambda) + \mathscr{A}(\mathbf{a}', \lambda) \mathscr{B}(\mathbf{b}, \lambda) - \mathscr{A}(\mathbf{a}', \lambda) \mathscr{B}(\mathbf{b}', \lambda) \right\} p(\lambda) \, d\lambda \leqslant +2. \quad (24)$$

Therefore, the equality (22) implies that the absolute value of the combination of expectation values is bounded by 2:

$$-2 \leqslant \mathcal{E}(\mathbf{a}, \mathbf{b}) + \mathcal{E}(\mathbf{a}, \mathbf{b}') + \mathcal{E}(\mathbf{a}', \mathbf{b}) - \mathcal{E}(\mathbf{a}', \mathbf{b}') \leqslant +2. \quad (25)$$

But since the bounds on (20) predicted by quantum mechanics and observed in experiments are $\pm 2\sqrt{2}$, Bell concludes that no local and realistic theory envisaged by Einstein can reproduce the statistical predictions of quantum mechanics. In particular, contextual hidden variable theories specified by (9) that respect the factorizability (17) are not viable.

In many derivations of (25) in the literature, factorized probabilities of observing binary measurement results are employed rather than measurement results themselves I have used in (17) in my derivation following Bell [1,17]. But employing probabilities would only manage to obfuscate the logical flaw in Bell's argument I intend to bring out here.

1.3.2. Converse Derivation of the Additivity (22) by Assuming (25)

Now, it is not difficult to demonstrate the *converse* of the above derivation in which the additivity of expectation values (22) is derived by assuming the stringent bounds of ± 2 on the sum (20). Employing (18), (20) can be written as

$$\int_{\mathscr{L}} \mathscr{A}(\mathbf{a}, \lambda) \mathscr{B}(\mathbf{b}, \lambda) \, p(\lambda) \, d\lambda + \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}, \lambda) \mathscr{B}(\mathbf{b}', \lambda) \, p(\lambda) \, d\lambda + \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}', \lambda) \mathscr{B}(\mathbf{b}, \lambda) \, p(\lambda) \, d\lambda - \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}', \lambda) \mathscr{B}(\mathbf{b}', \lambda) \, p(\lambda) \, d\lambda. \quad (26)$$

Since each product $\mathscr{A}(\mathbf{a}, \lambda) \mathscr{B}(\mathbf{b}, \lambda)$ in the above integrals is equal to ± 1, each of the four integrals is bounded by ± 1:

$$-1 \leqslant \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}, \lambda) \mathscr{B}(\mathbf{b}, \lambda) \, p(\lambda) \, d\lambda \leqslant +1. \quad (27)$$

Thus the sum of four integrals in (26) is bounded by ± 4, not ± 2. However, we started with (25), which contends that the sum of integrals in (26) is bounded by ± 2. But the only way to reduce the bounds on (26) from ± 4 to ± 2 without compromising the independence of the results $\mathscr{A}(\mathbf{a}, \lambda) \mathscr{B}(\mathbf{b}, \lambda)$, *etc.*, or of their averages appearing in the sum (26), or violating the rules of anti-derivatives, is by equating the sum of integrals in (26) to the following integral of the sum:

$$\int_{\mathscr{L}} \left\{ \mathscr{A}(\mathbf{a}, \lambda) \mathscr{B}(\mathbf{b}, \lambda) + \mathscr{A}(\mathbf{a}, \lambda) \mathscr{B}(\mathbf{b}', \lambda) + \mathscr{A}(\mathbf{a}', \lambda) \mathscr{B}(\mathbf{b}, \lambda) - \mathscr{A}(\mathbf{a}', \lambda) \mathscr{B}(\mathbf{b}', \lambda) \right\} p(\lambda) \, d\lambda. \quad (28)$$

To be sure, we can easily reduce the bounds on the sum in (26) from ± 4 to ± 2, for example, by setting the last two integrals in it equal to each other, with the minus sign between them. That would reduce (26) to a sum of only the first two integrals, which is then easily seen to be bounded by ± 2 using (27). But to achieve this, we have compromised the independence of at least the last two sub-ensemble averages, which is not permitted by the conditions adhered to in Bell's argument. Similarly, again as an example, we can set the results themselves in the last two integrals in (26) equal to each other: $\mathscr{A}(\mathbf{a}', \lambda) \mathscr{B}(\mathbf{b}, \lambda) = \mathscr{A}(\mathbf{a}', \lambda) \mathscr{B}(\mathbf{b}', \lambda)$. That would again reduce (26) to the sum of only the first two integrals, which is then easily seen to be bounded by ± 2 using (27). But we have again compromised the independence of the last two results to achieve this, which is not permitted. One can also easily reduce the bounds on (26) from ± 4 to ± 2 by violating some of the rules of anti-derivatives, or of statistical averages, *etc.*, as an extreme measure, but that again amounts to

forfeiting the game from the start. Thus, the only admissible means of reducing the bounds on (26) from ± 4 to ± 2 is by using the standard rules of anti-derivatives that allow us to equate the sum of integrals appearing in (26) to the integral (28) of the sum of eigenvalues, which, as we saw in (24), is bounded by ± 2.

We have thus derived the additivity of expectation values (22) by imposing the bounds of ± 2 on the Bell-CHSH sum (20) as our starting assumption. Therefore, given the previous derivation that led us to (25) by assuming (22) and the current derivation that led us to (22) by assuming (25), we have proved that the assumption (22) of the additivity of expectation values is tautologous to assuming the bounds of ± 2 on the Bell-CHSH sum (20) of expectation values.

1.4. Additivity of Expectation Values (22) is An Unjustified Assumption, Equivalent to the Thesis to be Proven

The key step that led us to the bounds of ± 2 on (20) that are more restrictive than $\pm 2\sqrt{2}$ is the step (22) of the linear additivity of expectation values. In what follows, I will demonstrate that this step is, in fact, an unjustified assumption, equivalent to the main thesis of the theorem to be proven, just as it is in von Neumann's now discredited theorem [9,18,25,28,29]. But this fact is obscured by the seemingly innocuous built-in linear additivity of integrals used in step (22). However, as we noted around (12) and will be further demonstrated in Section 1.6, the built-in linear additivity of integrals is physically meaningful only for simultaneously measurable or commuting observables [25,28]. It is, therefore, not legitimate to invoke it at step (22) *without proof*. Step (22) would be valid also in classical physics in which the value of a sum of observable quantities would be the same as the sum of the values each quantity would take separately, because, unlike in quantum mechanics, they would all be simultaneously measurable, yielding only sharp values. Perhaps for this reason it is usually not viewed as an assumption but mistaken for a benign mathematical step. It is also sometimes claimed to be necessitated by Einstein's requirement of realism [2]. But I will soon explain why it is a much overlooked unjustified assumption, and demonstrate in Section 1.6 that, far from being required by realism, the right-hand side of step (22), in fact, *contradicts* realism, which requires that *every* observable of a physical system, including any sums of observables, must be assigned a correct eigenvalue, quantifying one of its preexisting properties.

Moreover, realism has already been adequately accommodated by the very definition of the local functions $\mathscr{A}(\mathbf{a}, \lambda)$ and $\mathscr{B}(\mathbf{b}, \lambda)$ and their counterfactual juxtaposition on the left-hand side of (22), as contextually existing properties of the system. Evidently, while a result in only one of the four expectation values corresponding to a sub-experiment that appears on the left-hand side of (22) can be realized in a given run of a Bell-test experiment, the remaining three results appearing on that side are realizable at least counterfactually, thus fulfilling the requirement of realism [9]. Therefore, the requirement of realism does not necessitate the left-hand side of (22) to be equated with its right-hand side in the derivation of (25). Realism requires definite results $\mathscr{A}(\mathbf{a}, \lambda)\,\mathscr{B}(\mathbf{b}, \lambda)$ to exist as eigenvalues only counterfactually, *not all four at once*, as they are written on the right-hand side of (22). What is more, as we will soon see, realism implicit in the prescription (9) requires the quantity (23) to be a *correct* eigenvalue of the summed operator (39), but it is not.

On the other hand, given the assumption $p(\lambda \,|\, \mathbf{a}, \mathbf{b}) = p(\lambda)$ of statistical independence and the additivity property of anti-derivatives, mathematically the equality (22) follows at once because of the linearity built into the integrals, provided we adopt a double standard for additivity: we reject (22) for von Neumann's theorem as Bell did in [18], but accept it unreservedly for Bell's theorem [9,25]. The binary properties of the functions $\mathscr{A}(\mathbf{a}, \lambda)$ and $\mathscr{B}(\mathbf{b}, \lambda)$ then immediately lead us to the bounds of ± 2 on (20). But, as we saw above, assuming the bounds of ± 2 on (20) leads, conversely, to the assumption (22) of additivity of expectation values. Thus, assuming the additivity of expectation values (22) is mathematically equivalent to assuming the bounds of ± 2 on the Bell-CHSH sum (20). In other words, Bell's argument presented in Section 1.3 *assumes* its conclusion (25) in the guise of assumption (22), by

implicitly assuming that the expectation functions $(\Psi, \lambda \,|\, \Omega(c) \,|\, \Psi, \lambda)$ determining the eigenvalues $\omega(c, \lambda)$ of $\Omega(c)$ are *linear* [29]:

$$\widetilde{\omega}(\widetilde{c}, \lambda) = (\Psi, \lambda \,|\, \left[\sum_{i=1}^{4} \Omega_i(c_i)\right] \,|\, \Psi, \lambda) = \sum_{i=1}^{4} (\Psi, \lambda \,|\, \Omega_i(c_i) \,|\, \Psi, \lambda) = \sum_{i=1}^{4} \omega_i(c_i, \lambda). \qquad (29)$$

But, as explained by Bohm and Bub in [29] (see Appendix B below), this assumed linearity of $(\Psi, \lambda \,|\, \Omega(c) \,|\, \Psi, \lambda)$ is unreasonably restrictive for dispersion-free states $|\Psi, \lambda\rangle$, because the observables defined in (15) are not simultaneously measurable. However, it allows us to reduce the following correct relation within quantum mechanics *as well as* hidden variable theories,

$$\sum_{i=1}^{4} \left[\int_{\mathscr{L}} (\Psi, \lambda \,|\, \Omega_i(c_i) \,|\, \Psi, \lambda) \, p(\lambda) \, d\lambda \right] = \int_{\mathscr{L}} (\Psi, \lambda \,|\, \left[\sum_{i=1}^{4} \Omega_i(c_i)\right] \,|\, \Psi, \lambda) \, p(\lambda) \, d\lambda, \qquad (30)$$

to the relation

$$\sum_{i=1}^{4} \left[\int_{\mathscr{L}} (\Psi, \lambda \,|\, \Omega_i(c_i) \,|\, \Psi, \lambda) \, p(\lambda) \, d\lambda \right] = \int_{\mathscr{L}} \left[\sum_{i=1}^{4} (\Psi, \lambda \,|\, \Omega_i(c_i) \,|\, \Psi, \lambda) \right] p(\lambda) \, d\lambda, \qquad (31)$$

which is the same as assumption (22), albeit written in a more general notation. The equality (30), on the other hand, is equivalent to the quantum mechanical relation (32) discussed below, which can be verified using the prescription (9). The same equality (30) is also valid for hidden variable theories, because it does not make the mistake of relying on the linearity assumption (29). This can be verified also using (9) and the ansatz (3). Thus, the innocuous-looking linear additivity of integrals in assumption (22), while mathematically correct, is neither innocent nor physically reasonable.

It is not difficult to understand why appealing to the built-in linear additivity of anti-derivatives is not as innocent or physically reasonable as it may seem. In fact, for non-commuting observables that are not simultaneously measurable, justification of (22) or (31) by appealing to the built-in linear additivity of integrals leads to *incorrect equality between unequal physical quantities*. The reasons for this were recognized by Grete Hermann [25] some three decades before the formulation of Bell's theorem [1], as part of her insightful criticism of von Neumann's alleged theorem [3,28]. As she explained in [25], we are not concerned here with classical physics in which all observable quantities are simultaneously measurable yielding only sharp values, and therefore the value of a sum of observable quantities is nothing other than the sum of the values each of those quantities would separately take. Consequently, in classical physics, the averages of such values over individual initial states λ of the system can also be meaningfully added linearly, just as assumed in step (22) or (31), because there is no scope for any contradiction between the averages obtained by evaluating the left-hand side and the right-hand side of these equations. Therefore, in classical physics linear additivity of expectation values remains consistent with the built-in linear additivity of anti-derivatives. However, the same cannot be assumed without proof for the dispersion-free states $|\psi, \lambda\rangle$ of hidden variable theories, because, in that case, the values of the observable quantities are eigenvalues of the corresponding quantum mechanical operators dictated by the ansatz (3), and, as we noted above and toward the end of Section 1.1, the eigenvalue $\widetilde{\omega}(\widetilde{c}, \lambda)$ of the summed observable $\widetilde{\Omega}(\widetilde{c})$ is not equal to the sum $\sum_{i=1}^{n} \omega_i(c_i, \lambda)$ of the eigenvalues $\omega_i(c_i, \lambda)$ of $\Omega_i(c_i)$, unless the observables $\Omega_i(c_i)$ constituting the sum $\widetilde{\Omega}(\widetilde{c})$ are simultaneously measurable. *Thus, an important step in the proof of (25) is missing*. A necessary step that would prove the consistency of the built-in linear additivity of anti-derivatives with the non-additivity of expectation values for the *non-commuting* observables. In equation (42) of Section 1.6 below we will see the difference between the eigenvalue of the summed operator and the sum of individual eigenvalues explicitly. It will demonstrate how, in hidden variable theories equation (22) or (31) involving averages of eigenvalues ends up equating unequal averages

of physical quantities in general. It will thereby prove that, while valid in classical physics and for simultaneously measurable observables, equation (22) or (31) is *not* valid for hidden variable theories in general. Insisting otherwise thus amounts to *assuming* the validity of this equation *without proof*, despite the contrary evidence just presented [25]. That, in turn, amounts to assuming the very thesis to be proven — namely, the bounds of ± 2 on the Bell-CHSH sum (20). Consequently, the only correct meaning assignable to (22) or (25) is that it is valid only in classical physics and/or for commuting observables.

Sometimes assumption (22) is justified on statistical grounds. It is argued that the four sub-experiments appearing on the left-hand side of (22) with different experimental settings $\{\mathbf{a}, \mathbf{b}\}$, $\{\mathbf{a}, \mathbf{b}'\}$, *etc.* can be performed independently of each other, on possibly different occasions, and then the resulting averages are added together at a later time for statistical analysis. If the number of experimental runs for each pair of settings is sufficiently large, then, theoretically, the sum of the four averages appearing on the left-hand side of (22) are found not to exceed the bounds of ± 2, thus justifying the equality (22). This can be easily verified in numerical simulations (see Ref. [27] cited in [13]). However, this heuristic argument is not an analytical proof of the bounds. What it implicitly neglects to take into account by explicitly assuming that the four sub-experiments can be performed independently, is that the sub-experiments involve mutually exclusive pairs of settings such as $\{\mathbf{a}, \mathbf{b}\}$ and $\{\mathbf{a}, \mathbf{b}'\}$ in physical space, and thus involve non-commuting observables that cannot be measured simultaneously [9]. Unless the statistical analysis takes this physical fact into account, it cannot be claimed to have any relevance for the Bell-test experiments [16]. For ignoring this physical fact amounts to incorrectly assuming that the spin observables $\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}$, *etc.* are mutually commuting, and thus simultaneously measurable, for which assumption (22) is indeed valid, as demonstrated below in Section 1.6 (see the discussion around (45)). On the other hand, when the non-commutativity of the observables involved in the sub-experiments is taken into account in numerical simulations, the bounds on (20) turn out to be $\pm 2\sqrt{2}$, as shown in [10,11] and Ref. [27] cited in [13]. In other words, such a statistical argument is simply assumption (22) in disguise.

Another important point to recognize here is that the above derivation of the stringent bounds of ± 2 on (20) for a locally causal dispersion-free counterpart $|\Psi, \lambda\rangle$ of the quantum mechanical singlet state (13) must comply with the heuristics of the contextual hidden variable theories we discussed in Section 1.1. If it does not, then the bounds of ± 2 cannot be claimed to have any relevance for the viability of local hidden variable theories [24]. Therefore, as discussed in Section 1.1, in a contextual hidden variable theory all of the observables $\Omega_i(c_i)$ of any physical system, *including* their sum $\widetilde{\Omega}(\tilde{c}) = \sum_{i=1}^{n} \Omega_i(c_i)$, which also represents a physical quantity in the Hilbert space formulation of quantum mechanics [3] whether or not it is observed, must be assigned unique eigenvalues $\omega_i(c_i, \lambda)$ and $\widetilde{\omega}(\tilde{c}, \lambda)$, respectively, in the dispersion-free states $|\psi, \lambda\rangle$ of the system, regardless of whether these observables are simultaneously measurable. In particular, while the summed observable (39) discussed below is never observed in the Bell-test experiments, realism nevertheless requires it to be assigned a unique eigenvalue in accordance with the ansatz (3) and the prescription (9).

1.5. Additivity of Expectation Values is Respected by Quantum States

Now, within quantum mechanics, expectation values do add in analogy with the equality (22) assumed by Bell for local hidden variable theories [3,18]. In quantum mechanics, the statistical predictions of which any hidden variable theory is obliged to reproduce, the joint results $\mathscr{A}(\mathbf{a}, \lambda)\mathscr{B}(\mathbf{b}, \lambda)$ observed by Alice and Bob would be eigenvalues of the operators $\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}$, and the linearity in the rules of Hilbert space quantum mechanics ensures that these operators satisfy the additivity of expectation values. Thus, for any quantum state $|\psi\rangle$, the following equality holds:

$$\langle\psi|\,\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}\,|\psi\rangle + \langle\psi|\,\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}'\,|\psi\rangle + \langle\psi|\,\sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b}\,|\psi\rangle - \langle\psi|\,\sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b}'\,|\psi\rangle$$

$$= \langle\psi|\,\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b} + \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}' + \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b} - \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b}'\,|\psi\rangle. \qquad (32)$$

The linearity of this equation has been achieved by promoting observable quantities to Hermitian operators [3,25]. Any local hidden variable theory is then obliged to reproduce the predictions of both sides of this equation. Comparing equations (22) and (32), the equality between the two sides of (22) seems reasonable, even physically. Furthermore, since the condition (18) for any hidden variable theory obliges us to set the four terms on the left-hand side of (32) as

$$\langle \Psi | \, \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b} \, | \Psi \rangle = \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}, \lambda) \, \mathscr{B}(\mathbf{b}, \lambda) \, p(\lambda) \, d\lambda \,, \tag{33}$$

$$\langle \Psi | \, \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}' \, | \Psi \rangle = \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}, \lambda) \, \mathscr{B}(\mathbf{b}', \lambda) \, p(\lambda) \, d\lambda \,, \tag{34}$$

$$\langle \Psi | \, \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b} \, | \Psi \rangle = \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}', \lambda) \, \mathscr{B}(\mathbf{b}, \lambda) \, p(\lambda) \, d\lambda \,, \tag{35}$$

$$\text{and } \ \langle \Psi | \, \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b}' \, | \Psi \rangle = \int_{\mathscr{L}} \mathscr{A}(\mathbf{a}', \lambda) \, \mathscr{B}(\mathbf{b}', \lambda) \, p(\lambda) \, d\lambda \,, \tag{36}$$

it may seem reasonable that, given the quantum mechanical equality (32), any hidden variable theory should satisfy

$$\begin{aligned} \langle \Psi | \, \widetilde{\Omega}(\tilde{c}) \, | \Psi \rangle &= \langle \Psi | \, \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b} + \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}' + \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b} - \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b}' \, | \Psi \rangle \\ &= \int_{\mathscr{L}} \Big\{ \mathscr{A}(\mathbf{a}, \lambda) \, \mathscr{B}(\mathbf{b}, \lambda) + \mathscr{A}(\mathbf{a}, \lambda) \, \mathscr{B}(\mathbf{b}', \lambda) + \mathscr{A}(\mathbf{a}', \lambda) \, \mathscr{B}(\mathbf{b}, \lambda) - \mathscr{A}(\mathbf{a}', \lambda) \, \mathscr{B}(\mathbf{b}', \lambda) \Big\} \, p(\lambda) \, d\lambda \,, \end{aligned} \tag{37}$$

adhering to the prescription (9), which would then justify equality (22). Since hidden variable theories are required to satisfy the prescription (9), should not they also reproduce equation (37)? The answer to this is not straightforward.

1.6. Additivity of Expectation Values Does Not Hold for Dispersion-Free States

The problem with equation (37) is that, while the joint results $\mathscr{A}(\mathbf{a}, \lambda)\mathscr{B}(\mathbf{b}, \lambda)$, *etc.* appearing on the left-hand side of equation (22) are possible eigenvalues of the products of spin operators $\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}$, *etc.*, their summation

$$\mathscr{A}(\mathbf{a}, \lambda) \, \mathscr{B}(\mathbf{b}, \lambda) + \mathscr{A}(\mathbf{a}, \lambda) \, \mathscr{B}(\mathbf{b}', \lambda) + \mathscr{A}(\mathbf{a}', \lambda) \, \mathscr{B}(\mathbf{b}, \lambda) - \mathscr{A}(\mathbf{a}', \lambda) \, \mathscr{B}(\mathbf{b}', \lambda) \tag{38}$$

appearing as the integrand on the right-hand side of equation (37) or (22) is *not* an eigenvalue of the summed operator

$$\widetilde{\Omega}(\tilde{c}) = \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b} + \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}' + \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b} - \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b}', \tag{39}$$

because the spin operators $\sigma_1 \cdot \mathbf{a}$ and $\sigma_1 \cdot \mathbf{a}'$, *etc.*, and therefore $\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}$, *etc.*, do not commute with each other:

$$\begin{aligned} \big[\, \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}, \ \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}' \, \big] &= 2\,\sigma \cdot \big\{ (\mathbf{a} \times \mathbf{b}') \times (\mathbf{a} \times \mathbf{b}) \big\} \\ &\neq 0 \ \text{if} \ \mathbf{b}' \neq \mathbf{b} \neq \mathbf{a}. \end{aligned} \tag{40}$$

Consequently, equation (37) would hold within any hidden variable theory *only if* the operators $\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}$, *etc.* were commuting operators. As we discussed, this is well known from the famous criticisms of von Neumann's theorem against hidden variable theories [9,18,25,28]. While the equality (22) of the sum of expectation values with the expectation value of the sum is respected in quantum mechanics, it does not hold for hidden variable theories [18]. Nor does local realism necessitate the linear additivity (29) of eigenvalues for individual dispersion-free states $| \Psi, \lambda \rangle$.

This problem, however, suggests its own resolution. We can work out the correct eigenvalue $\widetilde{\omega}(\tilde{c}, \lambda)$ of the summed operator (39), at least formally, as I have worked out in Appendix C below. The correct version of equation (37) is then

$$\langle \Psi | \, \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b} + \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}' + \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b} - \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b}' \, | \Psi \rangle = \int_{\mathscr{L}} \widetilde{\omega}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}', \lambda) \, p(\lambda) \, d\lambda \,, \tag{41}$$

where

$$\widetilde{\omega} = \pm \sqrt{\left\{ \mathscr{A}(\mathbf{a}, \lambda)\, \mathscr{B}(\mathbf{b}, \lambda) + \mathscr{A}(\mathbf{a}, \lambda)\, \mathscr{B}(\mathbf{b}', \lambda) + \mathscr{A}(\mathbf{a}', \lambda)\, \mathscr{B}(\mathbf{b}, \lambda) - \mathscr{A}(\mathbf{a}', \lambda)\, \mathscr{B}(\mathbf{b}', \lambda) \right\}^2 + (\Psi, \lambda \,|\, \widetilde{\Theta} \,|\, \Psi, \lambda)} \tag{42}$$

is the correct eigenvalue of the summed operator (39), with its non-commuting part separated out as the operator

$$\widetilde{\Theta}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}') = 2\, \sigma \cdot \mathbf{n}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}') \,. \tag{43}$$

Here $(\Psi, \lambda \,|\, \widetilde{\Theta} \,|\, \Psi, \lambda) \neq 0$ in general, because the vector $\mathbf{n}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}')$ does not vanish in general. It works out to be

$$\mathbf{n}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}') = \{ \, (\mathbf{a} \times \mathbf{b}') \times (\mathbf{a} \times \mathbf{b}) + (\mathbf{a}' \times \mathbf{b}) \times (\mathbf{a} \times \mathbf{b}) + (\mathbf{a}' \times \mathbf{b}) \times (\mathbf{a} \times \mathbf{b}')$$
$$- (\mathbf{a}' \times \mathbf{b}') \times (\mathbf{a} \times \mathbf{b}) - (\mathbf{a}' \times \mathbf{b}') \times (\mathbf{a}' \times \mathbf{b}) - (\mathbf{a}' \times \mathbf{b}') \times (\mathbf{a} \times \mathbf{b}') \, \}. \tag{44}$$

The details of how this separation is accomplished using (40) can be found in Appendix C below. From (42), it is now easy to appreciate that the additivity of expectation values (22) assumed by Bell can hold only if the expectation value $(\Psi, \lambda \,|\, \widetilde{\Theta} \,|\, \Psi, \lambda) = \pm 2 \|\mathbf{n}\|$ of the non-commuting part within the eigenvalue $\widetilde{\omega}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}', \lambda)$ of the summed operator (39) is zero. But that is possible only if the operators $\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}$, *etc.* constituting the sum (39) commute with each other. In general, if the operators $\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}$, *etc.* in (39) do not commute with each other, then we would have

$$\widetilde{\omega}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}', \lambda) \neq \mathscr{A}(\mathbf{a}, \lambda)\, \mathscr{B}(\mathbf{b}, \lambda) + \mathscr{A}(\mathbf{a}, \lambda)\, \mathscr{B}(\mathbf{b}', \lambda) + \mathscr{A}(\mathbf{a}', \lambda)\, \mathscr{B}(\mathbf{b}, \lambda) - \mathscr{A}(\mathbf{a}', \lambda)\, \mathscr{B}(\mathbf{b}', \lambda). \tag{45}$$

But the operators $\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}$, *etc.* indeed do not commute with each other, because the pairs of directions $\{\mathbf{a}, \mathbf{a}'\}$, *etc.* in (39) are mutually exclusive directions in $\mathbb{R}^3$. Therefore, the additivity of expectation values assumed at step (22) in the derivation of (25) is unjustifiable. Far from being necessitated by realism, it actually contradicts realism.

Since three of the four results appearing in the expression (38) can be realized only counterfactually, their summation in (38) cannot be realized *even* counterfactually [9]. Thus, in addition to not being a correct eigenvalue of the summed operator (39) as required by the prescription (9) for hidden variable theories, the quantity appearing in (38) is, in fact, an entirely fictitious quantity, with no counterpart in any possible world, apart from in the trivial case when all observables are commutative. By contrast, the correct eigenvalue (42) of the summed operator (39) can be realized at least counterfactually because it is a genuine eigenvalue of that operator, thereby satisfying the requirement of realism correctly, in accordance with the prescription (9) for hidden variable theories. Using (42), all five of the observables appearing on both sides of the quantum mechanical equation (32) can be assigned unique and correct eigenvalues [9].

Once this oversight is ameliorated, it is not difficult to show that the conclusion of Bell's theorem no longer follows. For then, using the correct eigenvalue (42) of (39) instead of (38) on the right-hand side of (22), we have the equation

$$\mathcal{E}(\mathbf{a}, \mathbf{b}) + \mathcal{E}(\mathbf{a}, \mathbf{b}') + \mathcal{E}(\mathbf{a}', \mathbf{b}) - \mathcal{E}(\mathbf{a}', \mathbf{b}') = \int_{\mathscr{L}} \widetilde{\omega}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}', \lambda) \, p(\lambda) \, d\lambda \tag{46}$$

instead of (22), which implements local realism correctly on both of its sides, as required by the prescription (9) we discussed in Section 1.1. This equation (46) is thus the correct dispersion-free counterpart of the equivalence (32) for the quantum mechanical expectation values [9]. It can reduce to Bell's assumption (22) only when the expectation value $(\Psi, \lambda \,|\, \widetilde{\Theta} \,|\, \Psi, \lambda)$ of the non-commuting part within the eigenvalue $\widetilde{\omega}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}', \lambda)$ of the summed operator (39) happens to be vanishing. It thus expresses the correct relationship (30) among the expectation values for the singlet state (13) in the local hidden variable framework considered by Bell [1]. Recall again from the end of Section 1.1 that the quantum mechanical relation (32) is an unusual property of the quantum states $|\psi\rangle$. As Bell stressed in [18], "[t]here is no reason to demand it individually of the hypothetical dispersion free states, whose function it is to reproduce the *measurable* peculiarities of quantum mechanics *when averaged over*." Moreover, in Section V of [9] I have demonstrated that the bounds on the right-hand side of (46) are $\pm 2\sqrt{2}$ instead of ± 2. An alternative derivation of these bounds follows from the magnitude $||\mathbf{n}||$ of the vector defined in (44), which, as proved in Appendix D below, is bounded by 2, and therefore the eigenvalue $\pm 2\,||\mathbf{n}||$ of the operator (43) obtained as its expectation value $(\Psi, \lambda \,|\, \widetilde{\Theta} \,|\, \Psi, \lambda)$ is bounded by ± 4, giving

$$-4 \;\leqslant\; (\Psi, \lambda \,|\, \widetilde{\Theta}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}') \,|\, \Psi, \lambda) \;\leqslant\; +4\,. \tag{47}$$

Substituting these into (42), together with the bounds of ± 2 we worked out before on the commuting part (38), gives

$$-2\sqrt{2} \;\leqslant\; \widetilde{\omega}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}', \lambda) \;\leqslant\; +2\sqrt{2}\,, \tag{48}$$

which is constrained to be real despite the square root in the expression (42) because the operator (39) is Hermitian. Consequently, we obtain the following Tsirel'son's bounds in the dispersion-free state, on the right-hand side of (46):

$$-2\sqrt{2} \;\leqslant\; \int_{\mathscr{L}} \widetilde{\omega}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}', \lambda)\, p(\lambda)\, d\lambda \;\leqslant\; +2\sqrt{2}\,. \tag{49}$$

Given the correct relation (46) between expectation values instead of the flawed assumption (22), we thus arrive at

$$-2\sqrt{2} \;\leqslant\; \mathcal{E}(\mathbf{a}, \mathbf{b}) + \mathcal{E}(\mathbf{a}, \mathbf{b}') + \mathcal{E}(\mathbf{a}', \mathbf{b}) - \mathcal{E}(\mathbf{a}', \mathbf{b}') \;\leqslant\; +2\sqrt{2}\,. \tag{50}$$

Since the bounds of $\pm 2\sqrt{2}$ we have derived on the Bell-CHSH sum of expectation values are the same as those predicted by quantum mechanics and observed in the Bell-test experiments, the conclusion of Bell's theorem is mitigated. What is ruled out by these experiments is not local realism but the assumption of the additivity of expectation values, which does not hold for non-commuting observables in dispersion-free states of any hidden variable theories to begin with.

It is also instructive to note that the intermediate bounds $\pm 2\sqrt{2}$ on the Bell-CHSH sum (20), instead of the extreme bounds ± 2 or ± 4, follow in the above derivation of (50) as a consequence of the geometry of physical space [5,6]. Thus, what is brought out in it is the oversight of the non-commutative or Clifford-algebraic attributes of the physical space in Bell's derivation of the bounds ± 2 in (25). Indeed, it is evident from Appendix D below that the geometry of physical space imposes the bounds $0 \leqslant ||\mathbf{n}|| \leqslant 2$ on the magnitude of the vector (44), which, in turn, lead us to the bounds $\pm 2\sqrt{2}$ in (50). This is in sharp contrast with the traditional view of these bounds as due to non-local influences, stemming from a failure of the locality condition (17). But in the derivation of (50) above, the condition (17) is strictly respected. Therefore, the strength of the bounds $\pm 2\sqrt{2}$ in (50) is a consequence — *not* of non-locality or non-reality, but of the geometry of physical space [5,6]. Non-locality or non-reality is necessitated only if one erroneously insists on linear additivity (22) of eigenvalues of non-commuting observables for each *individual* dispersion-free state $|\Psi, \lambda\rangle$.

1.7. Conclusion: Bell's Theorem Assumes its Conclusion (Petitio Principii)

Let me reiterate the main points discussed above. Together, they demonstrate that Bell's theorem begs the question.

(1) The first point is that the derivation in Section 1.3 of the bounds of ± 2 on (20) for the dispersion-free counterpart $|\Psi, \lambda\rangle$ of the singlet state (13) must comply with the heuristics of the contextual hidden variable theories discussed in Section 1.1. Otherwise, the stringent bounds of ± 2 cannot be claimed to have any relevance for hidden variable theories. This requires compliance with the prescription (9) that equates the quantum mechanical expectation values with their hidden variable counterparts for *all* observables, including any sums of observables, pertaining to the singlet system.

(2) The most charitable view of the equality (22) is that it is an *assumption*, over and above those of locality, realism, and all other auxiliary assumptions required for deriving the inequalities (25), because *it is valid only for commuting observables*. Far from being required by realism, it contradicts realism, because it fails to assign the correct eigenvalue (42) to the summed observable (39) as its realistic counterpart, as required by the prescription (9). Realism requires that all observables, including their sums, must be assigned unique eigenvalues, regardless of whether they are observed.

(3) Expectation values in dispersion-free states of hidden variable theories do not add linearly for observables that are not simultaneously measurable. And yet, Bell assumed linear additivity (22) within a local hidden variable model. Conversely, in the light of the heuristics of contextual hidden variable theories we discussed in Section 1.1, assuming (22) is equivalent to assuming that the spin observables $\sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}$, *etc.* commute with each other, but they do not.

(4) When the correct eigenvalue (42) is assigned to the summed operator (39) replacing the incorrect step (22), the bounds on Bell-CHSH sum (20) work out to be $\pm 2\sqrt{2}$ instead of ± 2, thus mitigating the conclusion of Bell's theorem.

(5) As we proved in Section 1.3, the assumption (22) of the additivity of expectation values is equivalent to assuming the strong bounds of ± 2 on Bell-CHSH sum (20) of expectation values. In other words, (22) and (25) are tautologous.

(6) For observables that are not simultaneously measurable or commuting, the built-in linear additivity of integrals in step (22) or (31) leads to incorrect equality between averages of unequal physical quantities. Therefore, the view that this step is a harmless mathematical step is mistaken. It is, in fact, an unjustified assumption that is equivalent to the very thesis of the theorem to be proven, and is valid only in classical physics and/or for commuting observables.

The first four points above invalidate assumption (22), and thus inequalities (25) on physical grounds,   and   the   last two demonstrate that Bell's theorem assumes its conclusion in a different guise, and is thus invalid on logical grounds.

Bell's theorem is useful, however, for ruling out classical local theories. By relying on the assumption (22), which is valid for classical theories [25], it proves that no classical local theory can reproduce all of the predictions of quantum mechanics. But no serious hidden variable theories I am aware of have ever advocated returning to classical physics [19].

In this paper, I have focused on a formal or logical critique of Bell's theorem. Elsewhere [10,14,16], I have developed a comprehensive local-realistic framework for understanding quantum correlations in terms of the geometry of the spatial part of one of the well-known solutions of Einstein's field equations of general relativity — namely, that of a quaternionic 3-sphere — taken as a physical space within which we are confined to perform Bell-test experiments. This framework is based on Clifford algebra and thus explicitly takes the non-commutativity of observables into account. It thus shows, constructively, that contextually local hidden variable theories are not ruled

out by Bell-test experiments. Since, as we discussed in Section 1.2, the formal proof of Bell's theorem is based on the entangled singlet state (13), in [5,6,8,11–13,15] I have reproduced the correlations predicted by (13) as a special case within the local-realistic framework proposed in [10,14,16]. I especially recommend the calculations presented in [8] and [15], which also discuss a macroscopic experiment that would be able to falsify the 3-sphere hypothesis I have proposed in these publications.

Appendix

Appendix A Dynamical Equivalence of Quantum Mechanical Description and Einstein's Description

To establish the equivalence between two dynamics exhibited in Eq. (8), let us begin with the Ehrenfest equation,

$$\frac{d}{dt}\langle\psi\,|\,\Omega\,|\,\psi\rangle \;=\; \frac{1}{i\hbar}\langle\psi\,|\,[\,\Omega,\,H\,]\,|\,\psi\rangle + \langle\psi\,|\,\frac{\partial\Omega}{\partial t}\,|\,\psi\rangle, \tag{A1}$$

where H is a Hamiltonian operator. Using prescription (6) for the complete description, this equation can be written as

$$\frac{d}{dt}\int_{\mathscr{L}}(\,\psi,\lambda\,|\,\Omega\,|\,\psi,\lambda\,)\,p(\lambda)\,d\lambda \;=\; \frac{1}{i\hbar}\int_{\mathscr{L}}(\,\psi,\lambda\,|\,[\,\Omega,\,H\,]\,|\,\psi,\lambda\,)\,p(\lambda)\,d\lambda + \int_{\mathscr{L}}(\,\psi,\lambda\,|\,\frac{\partial\Omega}{\partial t}\,|\,\psi,\lambda\,)\,p(\lambda)\,d\lambda. \tag{A2}$$

Since in the ansatz (3) we have implicitly assumed that the hidden variables λ do not depend on time explicitly, the Leibniz integral rule for differentiation allows us to reduce the first term on the left-hand side of the above equation to

$$\frac{d}{dt}\int_{\mathscr{L}}(\,\psi,\lambda\,|\,\Omega\,|\,\psi,\lambda\,)\,p(\lambda)\,d\lambda \;=\; \frac{d}{dt}\int_{\mathscr{L}}\omega(\lambda)\,p(\lambda)\,d\lambda \;=\; \int_{\mathscr{L}}\left[\frac{d}{dt}\,\omega(\lambda)\right]p(\lambda)\,d\lambda. \tag{A3}$$

Next, we recall the Dirac rule for canonical quantization, which promotes the Poisson brackets to operators and sets

$$[\,\Omega,\,H\,] = i\hbar\,\{\widehat{\omega,\,\mathscr{H}}\}, \tag{A4}$$

where ω is an eigenvalue of the observable Ω and the Hamiltonian function $\mathscr{H}$ is an eigenvalue of the corresponding operator H. Then the Poisson bracket $\{\omega,\,\mathscr{H}\}$ is an eigenvalue of the operator $\{\widehat{\omega,\,\mathscr{H}}\}$. Now it is well known that the Dirac rule is not unique in general when promoting classical variables to quantum operators. It is unique, however, for the reverse process of reducing the quantum operators to classical variables. Given any two operators Ω_1 and Ω_2 with eigenvalues ω_1 and ω_2, their commutator $[\,\Omega_1,\,\Omega_2\,]$ uniquely reduces to the Poisson bracket $\{\omega_1,\,\omega_2\}$ in correspondence limit. Consequently, using (A4) and the ansatz (3), we can transform the first term on the right-hand side of (A2) to

$$\frac{1}{i\hbar}\int_{\mathscr{L}}(\,\psi,\lambda\,|\,[\,\Omega,\,H\,]\,|\,\psi,\lambda\,)\,p(\lambda)\,d\lambda \;=\; \int_{\mathscr{L}}(\,\psi,\lambda\,|\,\{\widehat{\omega,\,\mathscr{H}}\}\,|\,\psi,\lambda\,)\,p(\lambda)\,d\lambda \;=\; \int_{\mathscr{L}}\{\omega(\lambda),\,\mathscr{H}(\lambda)\}\,p(\lambda)\,d\lambda, \tag{A5}$$

where a consistent application of ansatz (3) to all observables of the system requires that $\{\omega,\,\mathscr{H}\}(\lambda) = \{\omega(\lambda),\,\mathscr{H}(\lambda)\}$. Finally, using the ansatz (3) for evaluating the second term on the right-hand side of (A2), it is easy to prove that

$$\int_{\mathscr{L}}(\,\psi,\lambda\,|\,\frac{\partial\Omega}{\partial t}\,|\,\psi,\lambda\,)\,p(\lambda)\,d\lambda \;=\; \int_{\mathscr{L}}\frac{\partial\omega(\lambda)}{\partial t}\,p(\lambda)\,d\lambda. \tag{A6}$$

Putting all of the results we have obtained in (A2), (A3),(A5), and (A6), we can finally conclude using (A1) that

$$\left[\frac{d}{dt}\langle\psi\,|\,\Omega\,|\,\psi\rangle = \frac{1}{i\hbar}\langle\psi\,|\,[\Omega,H]\,|\,\psi\rangle + \langle\psi\,|\,\frac{\partial\Omega}{\partial t}\,|\,\psi\rangle\right] = \int_{\mathscr{L}}\left[\frac{d}{dt}\,\omega(\lambda) = \{\omega(\lambda),\,\mathscr{H}(\lambda)\} + \frac{\partial\,\omega(\lambda)}{\partial t}\right]p(\lambda)\,d\lambda\,.$$
(A7)

For specific values of λ with $p(\lambda) = 1$, the right-hand side of (A7) reduces to classical equations of motion. If, however, only a probability distribution $p(\lambda)$ of λ is known, then Ehrenfest's equation in quantum mechanics can be understood as an average of an ensemble of classical dynamics, with each of its members specifying a trajectory in phase space. All probabilistic statements about the system would then reduce to the incompleteness of our knowledge about the system.

Appendix B Demonstration of Linearity of the Expectation Function $(\,\psi,\,\lambda\,|\,\Omega(c)\,|\,\psi,\,\lambda\,)$ Assumed in (22)

For the demonstration of linearity of expectation function $(\,\psi,\,\lambda\,|\,\Omega(c)\,|\,\psi,\,\lambda\,)$ assumed in (22), let us follow similar demonstration by Bohm and Bub in [29] within the context of von Neumann's theorem. To that end, recall that in the Hilbert space model of quantum mechanics [3], every ensemble can be characterized by a statistical operator, say $\mathcal{W}(\leftarrow)$, of unit trace, representing the quantum state $|\psi\rangle$ of the system. In other words, there exists a linear Hermitian matrix $\mathcal{W}_{pq}(\psi)$ such that the expectation value in the state $|\psi\rangle$ of observable Ω can be expressed in the linear form

$$\langle\psi\,|\,\Omega\,|\,\psi\rangle = \mathrm{Tr}\,\{\mathcal{W}(\psi)\,\Omega\} = \sum_{pq}\mathcal{W}_{qp}(\psi)\,\Omega_{pq}\,.$$
(A8)

However, if dispersion in the measured values of Ω is assumed to be due to a distribution $p(\lambda)$ in the values of hidden variables λ over the ensemble of the system as in (6), then the expectation value of Ω should be the average over λ,

$$\langle\psi\,|\,\Omega\,|\,\psi\rangle = \int_{\mathscr{L}}(\,\psi,\,\lambda\,|\,\Omega\,|\,\psi,\,\lambda\,)\ p(\lambda)\,d\lambda = \sum_{pq}\int_{\mathscr{L}}\mathcal{W}_{qp}(\psi,\lambda)\,\Omega_{pq}\,p(\lambda)\,d\lambda\,,$$
(A9)

which amounts to replacing the matrix $\mathcal{W}_{qp}(\psi)$ with its average $\overline{\mathcal{W}}_{qp}(\psi)$ over the distribution of hidden variables λ:

$$\mathcal{W}_{qp}(\psi) \longrightarrow \overline{\mathcal{W}}_{qp}(\psi) = \int_{\mathscr{L}}\mathcal{W}_{qp}(\psi,\lambda)\ p(\lambda)\,d\lambda\,.$$
(A10)

For each particular value of λ, (A9) thus assumes expectation function to be of a linear form, similar to that in (A8):

$$(\,\psi,\,\lambda\,|\,\Omega\,|\,\psi,\,\lambda\,) = \mathrm{Tr}\,\{\mathcal{W}(\psi,\lambda)\,\Omega\} = \sum_{pq}\mathcal{W}_{qp}(\psi,\lambda)\,\Omega_{pq}\,.$$
(A11)

By now it is well-recognized that demanding this linear form of expectation function $(\,\psi,\,\lambda\,|\,\Omega\,|\,\psi,\,\lambda\,)$ is not justifiable for hidden variable theories [18,25,28,29]. Ironically, however, after criticizing von Neumann for retaining the linear form (A11), Bell also ended up making the same mistake for contextual hidden variable theories, by implicitly assuming

$$\int_{\mathscr{L}}(\,\psi,\,\lambda\,|\,\Omega(c)\,|\,\psi,\,\lambda\,)\ p(\lambda)\,d\lambda = \sum_{pq}\int_{\mathscr{L}}\mathcal{W}_{qp}(\psi,\lambda)\,\Omega_{pq}(c)\,p(\lambda)\,d\lambda$$
(A12)

in the equality (22). It is not difficult to see that (22) follows from the linearity (A11) of the function $(\,\psi,\,\lambda\,|\,\Omega(c)\,|\,\psi,\,\lambda\,)$:

$$\int_{\mathscr{L}}(\,\psi,\,\lambda\,|\,\Omega_1(c_1) + \Omega_2(c_2)\,|\,\psi,\,\lambda\,)\ p(\lambda)\,d\lambda = \sum_{pq}\int_{\mathscr{L}}\mathcal{W}_{qp}(\psi,\lambda)\,\{\Omega_1(c_1) + \Omega_2(c_2)\}_{pq}\,p(\lambda)\,d\lambda$$
(A13)

$$\int_{\mathscr{L}} \widetilde{\omega}(\tilde{c}, \lambda) \; p(\lambda) \, d\lambda \; = \; \sum_{pq} \int_{\mathscr{L}} \mathcal{W}_{qp}(\psi, \lambda) \; \{ (\Omega_1)_{pq}(c_1) + (\Omega_2)_{pq}(c_2) \} \; p(\lambda) \, d\lambda$$

(A14)

$$\int_{\mathscr{L}} \widetilde{\omega}(\tilde{c}, \lambda) \; p(\lambda) \, d\lambda \; = \; \sum_{pq} \int_{\mathscr{L}} \mathcal{W}_{qp}(\psi, \lambda) \; (\Omega_1)_{pq}(c_1) \; p(\lambda) \, d\lambda$$
$$+ \sum_{pq} \int_{\mathscr{L}} \mathcal{W}_{qp}(\psi, \lambda) \; (\Omega_2)_{pq}(c_2) \; p(\lambda) \, d\lambda$$

(A15)

$$\int_{\mathscr{L}} \widetilde{\omega}(\tilde{c}, \lambda) \; p(\lambda) \, d\lambda \; = \; \int_{\mathscr{L}} (\, \psi, \lambda \,|\, \Omega_1(c_1) \,|\, \psi, \lambda \,) \; p(\lambda) \, d\lambda \; + \int_{\mathscr{L}} (\, \psi, \lambda \,|\, \Omega_2(c_2) \,|\, \psi, \lambda \,) \; p(\lambda) \, d\lambda$$

(A16)

$$\int_{\mathscr{L}} \widetilde{\omega}(\tilde{c}, \lambda) \; p(\lambda) \, d\lambda \; = \; \int_{\mathscr{L}} \omega_1(c_1, \lambda) \; p(\lambda) \, d\lambda \; + \int_{\mathscr{L}} \omega_2(c_2, \lambda) \; p(\lambda) \, d\lambda \,,$$

(A17)

where only two terms of (22) is used for simplicity, and the ansatz (3) is used to give $(\, \psi, \lambda \,|\, \Omega_1(c_1) \,|\, \psi, \lambda \,) = \omega_1(c_1, \lambda)$, $(\, \psi, \lambda \,|\, \Omega_2(c_2) \,|\, \psi, \lambda \,) = \omega_2(c_2, \lambda)$, and $(\, \psi, \lambda \,|\, \Omega_1(c_1) + \Omega_2(c_2) \,|\, \psi, \lambda \,) = \widetilde{\omega}(\tilde{c}, \lambda)$. Then linearity of anti-derivatives gives

$$\int_{\mathscr{L}} \widetilde{\omega}(\tilde{c}, \lambda) \; p(\lambda) \, d\lambda \; = \; \int_{\mathscr{L}} \{ \omega_1(c_1, \lambda) + \omega_2(c_2, \lambda) \} \; p(\lambda) \, d\lambda \,,$$

(A18)

implying that the eigenvalue $\widetilde{\omega}(\tilde{c}, \lambda)$ of the sum $\Omega_1(c_1) + \Omega_2(c_2)$ of operators is equal to the sum of their eigenvalues:

$$\widetilde{\omega}(\tilde{c}, \lambda) = \omega_1(c_1, \lambda) + \omega_2(c_2, \lambda) \,,$$

(A19)

for *each* particular value of the hidden variables λ. But, as we noted in the last paragraph of Section 1.1, this equality among the eigenvalues is mistaken in general (recall the example $\sqrt{2} \neq 1 + 1$). It is valid only in classical physics, and for commuting observables. We have arrived at it by assuming the *linear* form of the expectation function $(\, \psi, \lambda \,|\, \Omega \,|\, \psi, \lambda \,)$ defined in (A11). Recognizing this difficulty, Bohm and Bub considered a more general, non-linear expectation function:

$$(\, \psi, \lambda \,|\, \Omega(c) \,|\, \psi, \lambda \,) = F(\psi, \lambda, \Omega_{pq}(c)) \neq \sum_{pq} \mathcal{W}_{qp}(\psi, \lambda) \, \Omega_{pq}(c) \,,$$

(A20)

where (at least for non-commuting observables) $F(\psi, \lambda, \Omega_{pq}(c))$ is a non-linear function of ψ, λ, and $\Omega_{pq}(c)$ such that

$$\widetilde{\omega}(\tilde{c}, \lambda) \neq \omega_1(c_1, \lambda) + \omega_2(c_2, \lambda) \,.$$

(A21)

Indeed, there is no reason why these eigenvalues should not be determined by some non-linear $F(\psi, \lambda, \Omega_{pq}(c))$, giving

$$\int_{\mathscr{L}} \omega_1(c_1, \lambda) \; p(\lambda) \, d\lambda \; + \int_{\mathscr{L}} \omega_2(c_2, \lambda) \; p(\lambda) \, d\lambda \; = \; \int_{\mathscr{L}} \widetilde{\omega}(\tilde{c}, \lambda) \; p(\lambda) \, d\lambda \; \neq \int_{\mathscr{L}} \{ \omega_1(c_1, \lambda) + \omega_2(c_2, \lambda) \} \; p(\lambda) \, d\lambda \,.$$

(A22)

As Bell stressed in [18], there is no reason for demanding (A19) of the dispersion-free state $|\, \psi, \lambda \,)$ for each value of λ.

Appendix C   Separating the Commuting and Non-Commuting Parts of the Summed Operator (39)

Before considering the specific operator (39), in this appendix let us prove that, in general, the eigenvalue of a sum $r\mathcal{R} + s\mathcal{S} + t\mathcal{T} + u\mathcal{U}$ of operators is not equal to the sum $r\mathscr{R} + s\mathscr{S} + t\mathscr{T} + u\mathscr{U}$ of the individual eigenvalues of the operators $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, and $\mathcal{U}$, unless these operators commute with each other. Here r, s, t, and u are real numbers. It is not difficult to prove this known fact by evaluating the square of the operator $\{ r\mathcal{R} + s\mathcal{S} + t\mathcal{T} + u\mathcal{U} \}$ as follows:

$$\{ r\mathcal{R} + s\mathcal{S} + t\mathcal{T} + u\mathcal{U} \} \{ r\mathcal{R} + s\mathcal{S} + t\mathcal{T} + u\mathcal{U} \} = r^2 \mathcal{R}^2 + rs\,\mathcal{R}\mathcal{S} + rt\,\mathcal{R}\mathcal{T} + ru\,\mathcal{R}\mathcal{U}$$

$$+ sr\,\mathcal{SR} + s^2\mathcal{S}^2 + st\,\mathcal{ST} + su\,\mathcal{SU}$$

$$+ tr\,\mathcal{TR} + ts\,\mathcal{TS} + t^2\mathcal{T}^2 + tu\,\mathcal{TU}$$

$$+ ur\,\mathcal{UR} + us\,\mathcal{US} + ut\,\mathcal{UT} + u^2\mathcal{U}^2. \tag{A23}$$

Now, assuming that the operators $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, and $\mathcal{U}$ do not commute in general, let us define the following operators:

$$\mathcal{L} := \mathcal{SR} - \mathcal{RS} \iff \mathcal{SR} = \mathcal{RS} + \mathcal{L}, \tag{A24}$$

$$\mathcal{M} := \mathcal{TR} - \mathcal{RT} \iff \mathcal{TR} = \mathcal{RT} + \mathcal{M}, \tag{A25}$$

$$\mathcal{N} := \mathcal{TS} - \mathcal{ST} \iff \mathcal{TS} = \mathcal{ST} + \mathcal{N}, \tag{A26}$$

$$\mathcal{O} := \mathcal{UR} - \mathcal{RU} \iff \mathcal{UR} = \mathcal{RU} + \mathcal{O}, \tag{A27}$$

$$\mathcal{P} := \mathcal{UT} - \mathcal{TU} \iff \mathcal{UT} = \mathcal{TU} + \mathcal{P}, \tag{A28}$$

$$\text{and} \quad \mathcal{Q} := \mathcal{US} - \mathcal{SU} \iff \mathcal{US} = \mathcal{SU} + \mathcal{Q}. \tag{A29}$$

These operators would be null operators with vanishing eigenvalues if the operators $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, and $\mathcal{U}$ did commute with each other. Using these relations for the operators $\mathcal{SR}, \mathcal{TR}, \mathcal{TS}, \mathcal{UR}, \mathcal{UT}$ and $\mathcal{US}$, equation (A23) can be simplified to

$$\{r\,\mathcal{R} + s\,\mathcal{S} + t\,\mathcal{T} + u\,\mathcal{U}\}\{r\,\mathcal{R} + s\,\mathcal{S} + t\,\mathcal{T} + u\,\mathcal{U}\} = r^2\mathcal{R}^2 + 2rs\,\mathcal{RS} + 2rt\,\mathcal{RT} + 2ru\,\mathcal{RU}$$

$$+ rs\,\mathcal{L} + s^2\mathcal{S}^2 + 2st\,\mathcal{ST} + 2su\,\mathcal{SU}$$

$$+ rt\,\mathcal{M} + st\,\mathcal{N} + t^2\mathcal{T}^2 + 2tu\,\mathcal{TU}$$

$$+ ru\,\mathcal{O} + su\,\mathcal{Q} + tu\,\mathcal{P} + u^2\mathcal{U}^2 \tag{A30}$$

$$= \{r\,\mathcal{R} + s\,\mathcal{S} + t\,\mathcal{T} + u\,\mathcal{U}\}_c^2 + \mathcal{Y}, \tag{A31}$$

where

$$\mathcal{Y} := rs\,\mathcal{L} + rt\,\mathcal{M} + st\,\mathcal{N} + ru\,\mathcal{O} + tu\,\mathcal{P} + su\,\mathcal{Q}. \tag{A32}$$

We have thus separated out the commuting part $\{r\,\mathcal{R} + s\,\mathcal{S} + t\,\mathcal{T} + u\,\mathcal{U}\}_c$ and the non-commuting part $\mathcal{Y}$ of the summed operator $\mathcal{X} := \{r\,\mathcal{R} + s\,\mathcal{S} + t\,\mathcal{T} + u\,\mathcal{U}\}$. Note that the operators $\mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}, \mathcal{P}$, and $\mathcal{Q}$ defined in (A24) to (A29) will not commute with each other in general unless their constituents $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, and $\mathcal{U}$ themselves are commuting. Next, we work out the eigenvalue $\mathscr{X}$ of the operator $\mathcal{X}$ in a normalized eigenstate $|\xi\rangle$ using the eigenvalue equations

$$\mathcal{X}\,|\xi\rangle = \mathscr{X}\,|\xi\rangle \tag{A33}$$

and

$$\mathcal{X}\,\mathcal{X}\,|\xi\rangle = \mathcal{X}\{\mathcal{X}\,|\xi\rangle\} = \mathcal{X}\{\mathscr{X}\,|\xi\rangle\} = \mathscr{X}\{\mathcal{X}\,|\xi\rangle\} = \mathscr{X}^2\,|\xi\rangle, \tag{A34}$$

in terms of the eigenvalues $\mathscr{R}$, $\mathscr{S}$, $\mathscr{T}$, and $\mathscr{U}$ of the operators $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, and $\mathcal{U}$ and the expectation value $\langle\xi|\mathcal{Y}|\xi\rangle$:

$$\mathscr{X} = \pm\sqrt{\langle\xi|\mathcal{X}\,\mathcal{X}|\xi\rangle} = \pm\sqrt{\langle\xi|\{r\,\mathcal{R} + s\,\mathcal{S} + t\,\mathcal{T} + u\,\mathcal{U}\}_c^2|\xi\rangle + \langle\xi|\mathcal{Y}|\xi\rangle}, \tag{A35}$$

where we have used (A31). But the eigenvalue of the commuting part $\{r\,\mathcal{R} + s\,\mathcal{S} + t\,\mathcal{T} + u\,\mathcal{U}\}_c$ of $\mathcal{X}$ is simply the linear sum $r\,\mathscr{R} + s\,\mathscr{S} + t\,\mathscr{T} + u\,\mathscr{U}$ of the eigenvalues of the operators $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, and $\mathcal{U}$. Consequently, using the equation analogous to (A34) for the square of the operator $\{r\,\mathcal{R} + s\,\mathcal{S} + t\,\mathcal{T} + u\,\mathcal{U}\}_c$ we can express the eigenvalue $\mathscr{X}$ of $\mathcal{X}$ as

$$\mathscr{X} = \pm\sqrt{\{r\,\mathscr{R} + s\,\mathscr{S} + t\,\mathscr{T} + u\,\mathscr{U}\}^2 + \langle\xi|\mathcal{Y}|\xi\rangle}. \tag{A36}$$

Now, because the operators $\mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}, \mathcal{P}$, and $\mathcal{Q}$ defined in (A24) to (A29) will not commute with each other in general if their constituent operators $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, and $\mathcal{U}$ are non-commuting, the state $|\xi\rangle$ will not be an eigenstate of the operator $\mathcal{Y}$ defined in (A32). Moreover, while a dispersion-free state $|\psi, \lambda\rangle$

would pick out one of the eigenvalues $\mathscr{Y}$ of y, it will not be equal to the linear sum of the corresponding eigenvalues $\mathscr{L}$, $\mathscr{M}$, $\mathscr{N}$, $\mathscr{O}$, $\mathscr{P}$, and $\mathscr{Q}$ in general,

$$\mathscr{Y} \neq rs\,\mathscr{L} + rt\,\mathscr{M} + st\,\mathscr{N} + ru\,\mathscr{O} + tu\,\mathscr{P} + su\,\mathscr{Q}, \tag{A37}$$

even if we assume that the operators x and y commute with each other so that $(\psi, \lambda \,|\, y \,|\, \psi, \lambda) = \mathscr{Y}$ is an eigenvalue of y. That is to say, just like the eigenvalue $\mathscr{X}$ of x, the eigenvalue $\mathscr{Y}$ of y is also a nonlinear function in general. On the other hand, because we wish to prove that the eigenvalue of the sum $r\mathscr{R} + s\mathscr{S} + t\mathscr{T} + u\mathscr{U}$ of the operators $\mathscr{R}$, $\mathscr{S}$, $\mathscr{T}$, and $\mathscr{U}$ is not equal to the sum $r\mathscr{R} + s\mathscr{S} + t\mathscr{T} + u\mathscr{U}$ of the individual eigenvalues of the operators $\mathscr{R}$, $\mathscr{S}$, $\mathscr{T}$, and $\mathscr{U}$ unless they commute with each other, we must make sure that the eigenvalue $\mathscr{Y}$ does not vanish for the unlikely case in which the operators $\mathscr{L}$, $\mathscr{M}$, $\mathscr{N}$, $\mathscr{O}$, $\mathscr{P}$, and $\mathscr{Q}$ commute with each other. But even in that unlikely case, we would have

$$\overline{\mathscr{Y}} = rs\,\mathscr{L} + rt\,\mathscr{M} + st\,\mathscr{N} + ru\,\mathscr{O} + tu\,\mathscr{P} + su\,\mathscr{Q} \tag{A38}$$

as eigenvalue of the operator y defined in (A32), and consequently the eigenvalue $\mathscr{X}$ in (A36) will at best reduce to

$$\mathscr{X} = \pm\sqrt{\left\{ r\mathscr{R} + s\mathscr{S} + t\mathscr{T} + u\mathscr{U} \right\}^2 + rs\,\mathscr{L} + rt\,\mathscr{M} + st\,\mathscr{N} + ru\,\mathscr{O} + tu\,\mathscr{P} + su\,\mathscr{Q}}. \tag{A39}$$

In other words, even in such an unlikely case $\mathscr{Y}$ will not vanish, and consequently the eigenvalue $\mathscr{X}$ will not reduce to

$$\overline{\mathscr{X}} = r\mathscr{R} + s\mathscr{S} + t\mathscr{T} + u\mathscr{U}. \tag{A40}$$

Consequently, unless $(\psi, \lambda \,|\, y(c) \,|\, \psi, \lambda) \equiv 0$, the expectation value of $x(c)$ equating the average of $\mathscr{X}(c, \lambda)$ will be

$$\langle \psi \,|\, \mathcal{X}(c) \,|\, \psi \rangle = \int_{\mathscr{L}} \mathscr{X}(c, \lambda)\, p(\lambda)\, d\lambda \tag{A41}$$

$$= \int_{\mathscr{L}} \left[\pm\sqrt{\left\{ r\,\mathscr{R}(c, \lambda) + s\,\mathscr{S}(c, \lambda) + t\,\mathscr{T}(c, \lambda) + u\,\mathscr{U}(c, \lambda) \right\}^2 + (\psi, \lambda \,|\, y(c) \,|\, \psi, \lambda)} \right] p(\lambda)\, d\lambda \tag{A42}$$

$$\neq \int_{\mathscr{L}} \left[\pm\sqrt{\left\{ r\,\mathscr{R}(c, \lambda) + s\,\mathscr{S}(c, \lambda) + t\,\mathscr{T}(c, \lambda) + u\,\mathscr{U}(c, \lambda) \right\}^2 + \overline{\mathscr{Y}}(c, \lambda)} \right] p(\lambda)\, d\lambda \text{ if } [\mathcal{X}, \mathcal{Y}] \neq 0 \tag{A43}$$

$$\neq \int_{\mathscr{L}} \pm\left\{ r\,\mathscr{R}(c, \lambda) + s\,\mathscr{S}(c, \lambda) + t\,\mathscr{T}(c, \lambda) + u\,\mathscr{U}(c, \lambda) \right\} p(\lambda)\, d\lambda \text{ if } \mathscr{L}, \mathscr{M}, \mathscr{N}, \mathscr{O}, \mathscr{P}, \mathscr{Q} \neq 0, \tag{A44}$$

where c indicates the contexts of experiments as discussed in Section 1.1. The above result confirms the inequality (45) we discussed in Section 1.6. Note that, because $\mathscr{X}(c, \lambda)$ and $\mathscr{Y}(c, \lambda)$ are highly *nonlinear* functions in general (recall, *e.g.*, that $\sqrt{x^2 \pm y^2} \neq \sqrt{x^2} \pm \sqrt{y^2}$), the inequality in (A44) can reduce to equality *if and only if* the operators $\mathscr{R}$, $\mathscr{S}$, $\mathscr{T}$, and $\mathscr{U}$ commute with each other. In that case, the operators $\mathscr{L}$, $\mathscr{M}$, $\mathscr{N}$, $\mathscr{O}$, $\mathscr{P}$, and $\mathscr{Q}$ defined in (A24) to (A29) will also commute with each other, as well as being null operators, with each of the eigenvalues $\mathscr{L}$, $\mathscr{M}$, $\mathscr{N}$, $\mathscr{O}$, $\mathscr{P}$, and $\mathscr{Q}$ reducing to zero. Consequently, in that case $(\psi, \lambda \,|\, y(c) \,|\, \psi, \lambda)$ will vanish identically and (A36) will reduce to (A40).

It is now straightforward to deduce the operator $\widetilde{\Theta}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}')$ specified in (43) using (40). For this purpose, we first note that for the Bell-CHSH sum (20) the real numbers $r = s = t = +1$ and $u = -1$, and therefore (A40) simplifies to

$$\overline{\mathscr{X}}(c, \lambda) = \mathscr{R}(c, \lambda) + \mathscr{S}(c, \lambda) + \mathscr{T}(c, \lambda) - \mathscr{U}(c, \lambda). \tag{A45}$$

This quantity is tacitly assumed in the derivation of Bell's theorem to be the eigenvalue of the summed operator (39), implying the following identifications:

$$\mathscr{A}(\mathbf{a}, \lambda)\,\mathscr{B}(\mathbf{b}, \lambda) \equiv \mathscr{R}(\mathbf{a}, \mathbf{b}, \lambda)$$
$$= \pm 1 \text{ is an eigenvalue of the observable } \mathcal{R}(\mathbf{a}, \mathbf{b}) \equiv \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}, \tag{A46}$$
$$\mathscr{A}(\mathbf{a}, \lambda)\,\mathscr{B}(\mathbf{b}', \lambda) \equiv \mathscr{S}(\mathbf{a}, \mathbf{b}', \lambda)$$

$$= \pm 1 \ \text{ is an eigenvalue of the observable } \ \mathcal{S}(\mathbf{a}, \mathbf{b}') \equiv \sigma_1 \cdot \mathbf{a} \otimes \sigma_2 \cdot \mathbf{b}', \tag{A47}$$

$$\mathscr{A}(\mathbf{a}', \lambda)\, \mathscr{B}(\mathbf{b}, \lambda) \equiv \mathscr{T}(\mathbf{a}', \mathbf{b}, \lambda)$$

$$= \pm 1 \ \text{ is an eigenvalue of the observable } \ \mathcal{T}(\mathbf{a}', \mathbf{b}) \equiv \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b}, \tag{A48}$$

$$\text{and} \quad \mathscr{A}(\mathbf{a}', \lambda)\, \mathscr{B}(\mathbf{b}', \lambda) \equiv \mathscr{U}(\mathbf{a}', \mathbf{b}', \lambda)$$

$$= \pm 1 \ \text{ is an eigenvalue of the observable } \ \mathcal{U}(\mathbf{a}', \mathbf{b}') \equiv \sigma_1 \cdot \mathbf{a}' \otimes \sigma_2 \cdot \mathbf{b}'. \tag{A49}$$

The non-commuting part of the operator (39) can therefore be identified using (A32) and the above identifications as

$$\widetilde{\Theta}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}') = \left\{ \mathcal{L} + \mathcal{M} + \mathcal{N} - \mathcal{O} - \mathcal{P} - \mathcal{Q} \right\}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}'), \tag{A50}$$

where the operators $\mathcal{L}$, $\mathcal{M}$, $\mathcal{N}$, $\mathcal{O}$, $\mathcal{P}$, and $\mathcal{Q}$ are defined in (A24) to (A29). The result is the operator specified in (43).

Appendix D Establishing Bounds on the Magnitude of the Vector n Defined in (44)

The vector $\mathbf{n}$ defined in (44) is a function of four unit vectors, $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$, and $\mathbf{b}'$, in $\mathbb{R}^3$, and involves various cross products among these vectors. Consequently, as the vectors $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$, and $\mathbf{b}'$ vary in their directions within $\mathbb{R}^3$ due to various choices made by Alice and Bob, the extremum values of the magnitude $||\mathbf{n}||$ is obtained by setting the vectors orthogonal to each other, with angles between them set to 90 or 270 degrees. However, in three dimensions that is possible only for three of the four vectors, so one of the four would have to be set either parallel or anti-parallel to one of the remaining three. Therefore, let us first choose to set $\mathbf{b}' = -\mathbf{b}$. Substituting this into (44) then gives $\mathbf{n} = \mathbf{0}$, and thus $||\mathbf{n}|| = 0$. We have thus found the lower bound on the magnitude $||\mathbf{n}||$. To determine the upper bound on $||\mathbf{n}||$, we set $\mathbf{a}' = -\mathbf{a}$ instead. Substituting this into (44) reduces the vector $\mathbf{n}$ to the following function of $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$ and $\mathbf{b}'$:

$$\mathbf{n} = 2\left\{ (\mathbf{a} \times \mathbf{b}') \times (\mathbf{a} \times \mathbf{b}) \right\}. \tag{A51}$$

Consequently, in this case, the magnitude of the vector $\mathbf{n}$ works out to be

$$||\mathbf{n}|| = 2\, ||(\mathbf{a} \times \mathbf{b}')||\, ||(\mathbf{a} \times \mathbf{b})||\, \sin\beta_{(\mathbf{a} \times \mathbf{b}'),(\mathbf{a} \times \mathbf{b})} \tag{A52}$$

$$= 2\left\{ ||\mathbf{a}||\, ||\mathbf{b}'||\, \sin\beta_{\mathbf{a},\mathbf{b}'} \right\}\left\{ ||\mathbf{a}||\, ||\mathbf{b}||\, \sin\beta_{\mathbf{a},\mathbf{b}} \right\}\left\{ \sin\beta_{(\mathbf{a} \times \mathbf{b}'),(\mathbf{a} \times \mathbf{b})} \right\}, \tag{A53}$$

where $\beta_{\mathbf{a},\mathbf{b}}$ is the angle between $\mathbf{a}$ and $\mathbf{b}$, *etc*. But since the vectors $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$, and $\mathbf{b}'$ are all unit vectors and we have set them orthogonal to each other (apart from $\mathbf{a}' = -\mathbf{a}$), we obtain $||\mathbf{n}|| = 2$ as the maximum possible value for the magnitude of $\mathbf{n}$. We have thus established the following bounds on the magnitude of the vector $\mathbf{n}$ as specified in (44):

$$0 \leqslant ||\mathbf{n}|| \leqslant 2. \tag{A54}$$

References

1. Bell JS. 1964 On the Einstein Podolsky Rosen paradox. *Physics.* **1**, 195–200. (doi:10.1103/PhysicsPhysiqueFizika.1.195)
2. Einstein A. 1948 Quantum mechanics and reality. *Dialectica.* **2**, 320–324. (doi:10.1111/j.1746-8361.1948.tb00704.x)
3. Von Neumann J. 1955 *Mathematical Foundations of Quantum Mechanics.* Princeton, NJ: Princeton University Press.
4. Coleman S., Mandula J. 1967 All possible symmetries of the S matrix. *Physical Review.* **159**, 1251–1256. (doi:10.1103/PhysRev.159.1251)
5. Christian J. 2019 Bell's theorem versus local realism in a quaternionic model of physical space. *IEEE Access.* **7**, 133388–133409. (doi:10.1109/ACCESS.2019.2941275)
6. Christian J. 2007 Disproof of Bell's theorem by Clifford algebra valued local variables. (doi:10.48550/arXiv.quant-ph/0703179)

7. Christian J. 2014 *Disproof of Bell's Theorem: Illuminating the Illusion of Entanglement*. Second Edition. Boca Raton, Florida: Brwonwalker Press.

8. Christian J. 2015 Macroscopic observability of spinorial sign changes under 2π rotations. *Int. J. Theor. Phys.* **54**, 20–46. (doi:10.1007/s10773-014-2412-2)

9. Christian J. 2017 Oversights in the respective theorems of von Neumann and Bell are homologous. (doi:10.48550/arXiv.1704.02876)

10. Christian J. 2018 Quantum correlations are weaved by the spinors of the Euclidean primitives. *R. Soc. Open Sci.* **5**, 180526. (doi:10.1098/rsos.180526)

11. Christian J. 2020 Dr. Bertlmann's socks in the quaternionic world of ambidextral reality. *IEEE Access.* **8**, 191028–191048. (doi:10.1109/ACCESS.2020.3031734)

12. Christian J. 2021 Reply to "Comment on 'Dr. Bertlmann's socks in the quaternionic world of ambidextral reality." *IEEE Access.* **9**, 72161–72171. (doi:10.1109/ACCESS.2021.3076449)

13. Christian J. 2022 Reply to "Comment on 'Bell's theorem versus local realism in a quaternionic model of physical space." *IEEE Access.* **10**, 14429–14439. (doi:10.1109/ACCESS.2022.3146421)

14. Christian J. 2022 Local origins of quantum correlations rooted in geometric algebra. (doi:10.48550/arXiv.2205.11372)

15. Christian J. 2022 Symmetric derivation of singlet correlations in a quaternionic 3-sphere model. (doi:10.48550/arXiv.2204.10288)

16. Christian J. 2022 Response to 'Comment on "Quantum correlations are weaved by the spinors of the Euclidean primitives."' *R. Soc. Open Sci.* **9**, 220147. (doi:10.1098/rsos.220147)

17. Clauser JF., Shimony A. 1978 Bell's theorem: Experimental tests and implications. *Rep. Prog. Phys.* **41**, 1881–1927. (doi:10.1088/0034-4885/41/12/002)

18. Bell JS. 1966 On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.* **38**, 447–452. (doi:10.1103/RevModPhys.38.447)

19. Gudder S. 1970 On hidden-variable theories. *Journal of Mathematical Physics*. **11**, 431–436. (doi:10.1063/1.1665156)

20. Robertson HP. 1929 The uncertainty principle. *Phys. Rev.* **34**, 163–164. (doi:10.1103/PhysRev.34.163)

21. Kochen S., Specker EP. 1967 The problem of hidden variables in quantum mechanics. *J. Math. Mech.* **17**, 59–87. (doi:10.1512/iumj.1968.17.17004)

22. Belinfante FJ. 1973 *A Survey of Hidden-Variables Theories*. New York, NY: Pergamon Press.

23. Gleason AM. 1957 Measures on the closed subspaces of a Hilbert space. *J. Math. Mech.* **6**, 885–893. (doi:10.1512/iumj.1957.6.56050)

24. Shimony A. 1984 Contextual hidden variables theories and Bell's inequalities. *Brit. J. Phil. Sci.* **35**, 25–45. (doi:10.1093/bjps/35.1.25)

25. Hermann G. 1935 Die naturphilosophischen Grundlagen der Quantenmechanik. *Abhandlungen der Fries'schen Schule* **6**, 75–152 (doi.org/10.1007/978-3-658-16241-2_12) — For English translation, see Chapter 15 of "Grete Hermann – Between Physics and Philosophy", Crull E. and Bacciagaluppi G., eds., Springer, 2016, 239–278 (doi:10.1007/978-94-024-0970-3)

26. Bohm D. 1951 *Quantum Theory*. Englewood Cliffs, NJ: Prentice-Hall. pp 614–623.

27. Clauser JF., Horne MA. 1974 Experimental consequences of objective local theories. *Phys. Rev. D*. **10**, 526–535. (doi:10.1103/PhysRevD.10.526)

28. Mermin ND., Schack R. 2018 Homer nodded: von Neumann's surprising oversight. *Found. Phys.* **48**, 1007–1020 (doi:10.1007/s10701-018-0197-5)

29. Bohm D., Bub J. 1966 A proposed solution of the measurement problem in quantum mechanics by a hidden variable theory. *Rev. Mod. Phys*. **38**, 453–469. (doi:10.1103/RevModPhys.38.453)