On the Separability of Physical Systems

Jon P. Jarrett

Abstract In the context of Bell-type experiments, two notions of "separability" emerge from the application of some simple considerations from information theory. The first of these applies to physical states construed as probability measures, while the second applies to states construed in terms of "elements of physical reality". The former is found to be logically equivalent to the "completeness" constraint (a.k.a. "outcome independence" or "factorizabilility") in Bell-type arguments. Moreover, it is found that there are theories that are separable in the first sense but which are empirically equivalent to no theory that is separable in the second sense. I offer some speculations about the significance of these and a few related results.

Bell's Theorem and the associated empirical tests of the Bell Inequalities are widely considered to reveal some strikingly non-classical features of our world. In the interest of trying to come to a fuller understanding of these features, I wish to propose some suggestions for how we might characterize the "separability" (or lack thereof) of systems in entangled states. For this purpose, I offer the following brief summary of the Bell milieu.[1, 2]

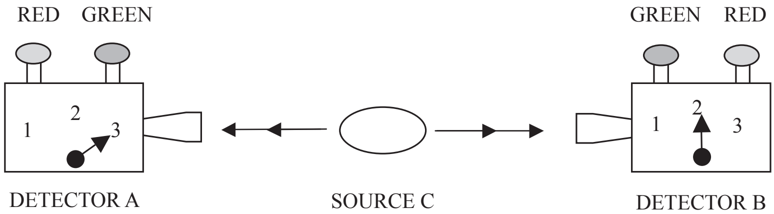
J.P. Jarrett
Department of Philosophy, University of Illinois at Chicago, IL, USA
e-mail: jarrett@uic.edu

[1] I find it congenial to address these issues in a framework consisting of an idealization of actual Bell-type experimental apparatuses, an idealization devised by N. David Mermin. I have dubbed Mermin's idealized experimental setup "the Mermin Contraption". It debuted (as simply "the device") in Mermin [1].

[2] This is intended as a handy reference for what is to follow. While most of this notation has become standard in the Bell literature, some of it (and some of my nomenclature) is rather idiosyncratic. See Jarrett [2] for further discussion.

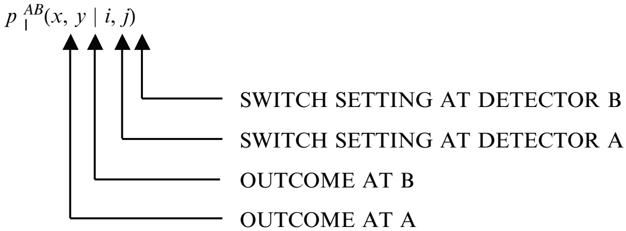
W.C. Myrvold and J. Christian (eds.), *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle,* The Western Ontario Series in Philosophy of Science 73,

The Mermin Contraption



COMPOSITE SYSTEM STATES: $\lambda \in \Lambda$
DETECTOR SETTINGS: 1, 2, 3
MEASUREMENT OUTCOMES: GREEN (+1), RED (−1)

Joint Probability Functions



Determinism

$$\forall \lambda, x, y, i, j,$$
$$p_\lambda^{AB}(x, y|i, j) \in \{0, 1\}$$

Theoretical Predictions

$$\forall \lambda, x, y, i, j,$$
$$P^{AB}(x, y|i, j) = \int_{\lambda \in \Lambda} \rho(\lambda) p_\lambda^{AB}(x, y|i, j) d\lambda,$$

where $\rho(\lambda)$ is the distribution of states.

The Mermin Contraption consists of three devices A, B and C. A and B are detectors of some sort, situated at diametrically opposed positions with respect to a source of particles, C. We make only some very limited assumptions regarding the nature of these devices and particles. Each detector can measure any one of three different things; the measurement outcomes consist of the illumination of either a red light or a green light at the detector. Distinct measurement-types correspond to distinct switch settings at a detector. Switch setting k at detector A corresponds to a measurement of the same type as switch setting k at detector B. A candidate theory for these experiments has to specify a set of states, Λ, for the two-particle system and must also provide an algorithm for computing joint probability functions of the form $p_{\lambda}(x,y \mid i,j)$. Here, λ is the state of the two particle system, x and y are the outcomes at site A and B, respectively; and i and j are the switch settings at site A and B, respectively. (A theory that satisfies DETERMINISM posits only states whose associated probability functions are restricted to $\{0,1\}$.) The testable predictions of the theory, then, are generated by summing these joint probability functions over all states, using a weighting function, $\rho(\lambda)$, the distribution of states, which the theory also must specify.

Marginal Probability Functions

$$\begin{aligned}&(\mathrm{A})\ p_{\lambda}^{A}(x|i,j) \equiv p_{\lambda}^{AB}(x,+1|i,j) + p_{\lambda}^{AB}(x,-1|i,j)\\ &(\mathrm{B})\ p_{\lambda}^{B}(y|i,j) \equiv p_{\lambda}^{AB}(+1,y|i,j) + p_{\lambda}^{AB}(-1,y|i,j)\end{aligned}$$

Locality

$$\begin{aligned}&\forall \lambda, x, i, i', j, j',\\ &\quad (\mathrm{A})\ p_{\lambda}^{A}(x|i,j) = p_{\lambda}^{A}\left(x|i,j'\right)\\ &\quad (\mathrm{B})\ p_{\lambda}^{B}(y|i,j) = p_{\lambda}^{B}\left(y|i',j\right)\end{aligned}$$

In terms of the joint probability functions, we can define marginal probabilities for A and B. I call *LOCALITY* the constraint that each of the marginals is independent of the switch setting at the distant site.[3] Relativity theory prohibits the superluminal transport of ma tter and energy. This relativistic "locality" constraint is by no means strictly equivalent to the foregoing LOCALITY. However, any theory that

[3] Abner Shimony prefers to call my LOCALITY condition *parameter independence*, and for good reason. This is a requirement that the probabilities at each detector site be independent of the switch setting (the "parameter") at the other detector site. "Locality" has been defined in various ways and taken to mean many different things; "parameter independence" however, is not only very descriptive, but also (in contrast to "locality") an expression whose meaning is univocal. For better or worse, I've chosen to retain this terminology, mainly out of deference to Einstein, Podolsky, and Rosen. I think that this constraint, while certainly not in any strict sense equivalent to the relativistic prohibition against superluminal signaling, does function in the Bell arguments in a manner that is very similar in spirit to the way that "locality," in the sense of EPR, was employed in their argument. Similar remarks apply to my use of the term "completeness."

violates LOCALITY faces at least a difficult reconciliation with relativity. (Note that quantum mechanics itself does satisfy LOCALITY.) Modulo a few related caveats, LOCALITY is a sufficient constraint from which to derive a Bell Inequality for theories that satisfy DETERMINISM. The results of actual experiments (in good agreement with the predictions of quantum mechanics) violate the Bell Inequality. Consequently, since relativity theory is so well confirmed independently, the experimental results are widely (though not universally) taken to provide rather direct evidence against DETERMINISM.[4]

Completeness

$$\begin{aligned}&\forall \lambda, x, y, i, j,\\ &p_\lambda^{AB}(x,y|i,j) = p_\lambda^A(x|i,j)\, p_\lambda^B(y|i,j)\end{aligned}$$

or, equivalently,

$$\begin{aligned}(\mathrm{A})\ p_\lambda^A(x|i,j,y) &\equiv \frac{p_\lambda^{AB}(x,y|i,j)}{p_\lambda^B(y|i,j)},\ \left(p_\lambda^B(y|i,j) \neq 0\right)\\ &= p_\lambda^A(x|i,j)\\ (\mathrm{B})\ p_\lambda^B(y|i,j,x) &\equiv \frac{p_\lambda^{AB}(x,y|i,j)}{p_\lambda^A(x|i,j)},\ \left(p_\lambda^A(x|i,j) \neq 0\right)\\ &= p_\lambda^B(y|i,j)\end{aligned}$$

Condition C

$$\begin{aligned}&\forall \lambda, i, j,\ \det C(\lambda|i,j) = 0,\\ &\text{where } C(\lambda|i,j) = \begin{bmatrix} p_\lambda^{AB}(+1,+1|i,j) & p_\lambda^{AB}(+1,-1|i,j)\\ p_\lambda^{AB}(-1,+1|i,j) & p_\lambda^{AB}(-1,-1|i,j)\end{bmatrix}\end{aligned}$$

CONDITION C $\Leftrightarrow$ COMPLETENESS

I call *COMPLETENESS* the requirement that for any specified state of the two-particle system and for any pair of measurement-types at the two detectors, the conditional probability of a specified outcome at either site given a specified outcome at the other site is just that (unconditioned) probability for getting the outcome at the first site.[5] To understand the sense in which this can be seen to be a kind of

[4] Most prominent among the dissenters are the advocates of a Bohmian mechanics. The Bohmians reject LOCALITY as a suitable constraint on physical theories.

[5] Shimony has called my COMPLETENESS constraint *outcome independence*, for it expresses the requirement that the probabilities at each detector site be independent of the outcome at the other detector site.

completeness condition, consider the following: an independence constraint of this sort is precisely that which should be satisfied if the information associated with the outcome at A(B) is redundant with respect to the probability assignments at B(A). Because of this redundancy, conditionalization on the outcome at A(B) cannot alter the probabilities assigned at B(A). The conditional probability of such and such at B(A) given the outcome at A(B) is the same as the unconditioned probability at B(A) because the (complete) state description already includes that information; there is nothing more that one can learn regarding what one might find at B(A) based on what one finds at A(B).

One might think of it this way: Our two particles of a given pair—call them "A" and "B"—interact with each other at the source. As a result of this interaction, B might leave its "particle footprint" or some of its "particle DNA" or whatever on A. An observation of A, then, could conceivably reveal some trace of information about B's properties, so that conditionalization on the measurement outcome at A could permit us to revise the probability for getting a specified outcome at B. However, it is precisely all such information that we suppose to be included antecedently in a genuinely complete state description. Thus, if our state description ***is*** complete, there should be no correlation between such pairs of measurement outcomes, and we would expect the joint probability functions to reduce to the product of the two marginals. For this reason, in some of the literature on Bell's Theorem, COMPLETENESS also gets called "factorizability", for it constrains the joint probabilities to factor into this product of the two marginals.[6]

COMPLETENESS demands that states posited by the theory afford a representation of phenomena in terms of the interactions of entities possessing physical properties in a suitably independent manner. State descriptions that are "complete" in this sense encode information in a manner that insures the "screening off" of any correlations between sets of measurement outcomes at the two wings of the Mermin Contraption. Any theory (whether deterministic or not) that posits entities that possess properties (whatever those properties may be) in the ways familiar to us from our experience at the classical level might plausibly be expected to satisfy this constraint. It has the look of nothing more than a requirement that state descriptions include all causally relevant information. (Note that quantum mechanics violates COMPLETENESS.)

CONDITION C deserves mention here, if only as a curiosity. I will address its possible significance later. The elements of the matrix $C(\lambda|i,j)$ are the joint probabilities assigned by λ to the four possible pairs of measurement outcomes, for specified switch settings i and j. CONDITION C is the requirement that the determinant of C vanish for all λ, i, and j. As can easily be verified by substitution of the relevant expressions for the marginal probabilities (and by applying the standard normalization constraint, according to which the probabilities for all possible pairs of outcomes sum to unity), COMPLETENESS and CONDITION C are logically equivalent.[7]

[6] The term "factorizability" was introduced by Arthur Fine, long-time nemesis of its referent.

[7] It is useful to appeal to this equivalence between COMPLETENESS and CONDITION C in order to see that theories that satisfy DETERMINISM also satisfy COMPLETENESS, while the

Strong Locality

$$\forall \lambda, x, y, i, j, i', j',$$
$$p_\lambda^{AB}(x, y|i, j) = p_\lambda^A(x|i, j')\, p_\lambda^B(y|i', j)$$

$$\begin{array}{c} \text{STRONG LOCALITY} \Longleftrightarrow \left\{ \begin{array}{c} \text{LOCALITY} \\ \& \\ \text{COMPLETENESS} \end{array} \right. \\ \Updownarrow \\ \text{GENERALIZED BELL INEQUALITIES} \end{array}$$

Finally, STRONG LOCALITY is logically equivalent to the conjunction of LOCALITY and COMPLETENESS. Theories that satisfy STRONG LOCALITY represent the composite system as a pair of component systems that are mutually independent in both the sense of LOCALITY and that of COMPLETENESS. Bell's Theorem and the related experimental results provide good evidence that no theory satisfying STRONG LOCALITY is empirically adequate, for it is this constraint from which the generalized Bell Inequalities are derived. These generalized Bell Inequalities, too, are violated by empirical data that are in excellent accord with the predictions of quantum mechanics. By a broad consensus, COMPLETENESS is assumed to be the culprit in all of this. However, it is crucial to bear in mind that theories that violate COMPLETENESS need not *ipso facto* have omitted anything whatsoever from their representation of physical states of affairs. That COMPLETENESS is satisfied by no acceptable empirically adequate theory cannot be regarded as some eliminable artifact of our currently best theory; rather, it appears to be a reflection of some highly non-classical feature of the world itself. Quantum mechanics appears to hook on to the world in some fashion that accurately mirrors this feature. In this respect, one might argue that Einstein, Podolsky, and Rosen (EPR) were quite right in concluding that quantum mechanics is "incomplete" in some important sense. Nevertheless, Bohr also appears to have been right, at least to the extent that "incompleteness" in this sense is demonstrably not a defect in a theory.[8] If (as I am assuming here) our theory must satisfy LOCALITY, and if that theory is to be empirically adequate, then it ***must*** violate COMPLETENESS. Again, this is ***not*** to say that any empirically adequate theory that satisfies LOCALITY fails to incorporate into its state descriptions all causally relevant factors; but it ***is*** to say that even if the state descriptions of such a theory ***do*** include that information (and, so, might still be called "complete" by that measure), those state descriptions nevertheless fail to screen off Bell-type correlations. This situation constitutes a much more radical departure from the classical worldview than is required by the mere rejection of DETERMINISM.

converse does not hold in general. If a theory satisfies DETERMINISM, then normalization demands that three of the four probabilities in matrix C have the value 0, while the other has the value 1. Thus, the determinant of C vanishes. However, it is easy to construct a 2×2 matrix of numbers that are in the interval [0,1) and which sum to 1 for which the determinant still vanishes.

[8] The classic papers here are Einstein, Podolsky, and Rosen [3] and Bohr [4].

The results of Bell-type experiments may be seen to impose constraints on our conception of causality and the strictures of relativity. COMPLETENESS itself may be regarded as a locality condition of a special sort, one that is closely related to the notion of "separability" (of the two component systems in Bell-type experiments). Below, I will introduce a notion of "separability" that turns out to be logically equivalent to COMPLETENESS in the previously described sense. However, as a candidate for a criterion of the individuation of physical systems, this is a comparatively weak separability constraint, certainly one that is weaker than any that would have drawn the support of EPR. Consequently, the mere incorporation into one's worldview of the "causal holism" implicit in the rejection of what I shall call *EPR-SEPARABILITY* is not by itself sufficient to move us into the realm of viable theories. Put in terms of how far we are compelled to retreat from the worldview of classical physics, giving up DETERMINISM is not enough; giving up EPR-SEPARABILITY is not enough; COMPLETENESS (this even weaker form of separability) must be abandoned as well.

If we accept the popular view that *the* lesson (or at least *one* of the lessons) that we must learn from Bell's Theorem is that apart from the empirical adequacy of quantum mechanics, apart from its future successes or ultimate demise, any theory that supersedes quantum mechanics is also going to have to violate COMPLETENESS,[9] then we must ask just what precisely it *means* to say of a theory that it is "incomplete". In particular, is there anything *more* to say about such theories beyond that which has already been said?[10]

Subsequent to the EPR paper, Einstein offered his own argument for the incompleteness of quantum mechanics.[11] Guided by considerations raised therein, Don Howard has suggested that we understand "separability" as a requirement that each of the putative subsystems possess its own distinct physical state in such a way that they fix the state of the composite system in precisely the manner specified by COMPLETENESS. For specified values of i and j, (marginal at A) represents the state of the A component system, (marginal at B) represents the state of the B component system, and (joint AB probability) represents the state of the composite system. COMPLETENESS is the requirement that the joint probability associated with λ, i, and j be fixed by the product of the two marginals. Thus, Howard recommends that COMPLETENESS be construed simply *as* separability.[12]

[9] I trust that even the Bohmians would allow that it is at least of some interest to consider what follows from this supposition.

[10] Whether or not Bell-type correlations should be deemed in need of further "explanation", so as to render them less "mysterious" is a question that impinges on some of the most fundamental issues in the philosophy of science. I do have views about the issues that bear on this question, but I will refrain from unleashing them here. An extremely stimulating discussion of these matters may be found in Cushing and McMullin [5]. (See, especially, Arthur Fine [6], Bas van Fraassen [7], and R. I. G. Hughes [8].)

[11] Historical scholarship suggests the possibility that Einstein did not even *see* the EPR paper until it appeared in print (see Fine [9], pp. 35–36), and that in any case he was not altogether pleased with the finished product. Einstein's own "incompleteness" argument first appeared in [10] and was later more fully elaborated in [11].

[12] See Howard [12] and [13].

Before acceding to Howard's recommendation, there are two important questions that need to be considered:

(i) Are probability functions **alone** properly regarded as providing an exhaustive characterization of physical states?

(ii) Even if physical states are to be understood in this manner, why ought a constraint deserving of the name "separability" *require* that the state of the composite system be determined by the states of the two component systems in the particular manner specified by COMPLETENESS? More specifically, why couldn't it be determined by those states in some other way and still, then, be considered as expressing an acceptable "separability" constraint?[13]

I will begin with question (ii). One way to formalize the talk of the "information" contained in state descriptions is in terms the Shannon information. This suggests the following definition:

Separability

$$\forall \lambda, i, j,$$
$$I\left(p_{\lambda}^{AB}|i,j\right) = I\left(p_{\lambda}^{A}|i,j\right) + I\left(p_{\lambda}^{B}|i,j\right),$$

where $I(p) \equiv \sum p \ln p$ is the Shannon information associated with probability function p, and the sum is over all values of the outcome variable(s) in p.

SEPARABILITY is just the requirement that the Shannon information associated with the joint probability function (representing the state of the composite system) equal the sum of the Shannon information associated with the marginal probability function at A (representing the state of the A component system)[14] and the Shannon information associated with the marginal probability function at B (representing the state of the B component system). This constraint, then, has a natural interpretation as one way, of perhaps many ways, one might choose to make precise the *cliché* that the whole is equal to the sum of the parts, and thereby serves as *a* "separability" constraint.

Interpreted in this manner, SEPARABILITY presupposes that states be construed as probability functions. Quantum mechanical states are in one-to-one correspondence with probability measures over the set of closed linear subspaces of a Hilbert space. So quantum mechanics is one example of a theory that effectively identifies states with probability functions. Classical mechanics can be regarded as a theory of that type as well, but in classical mechanics, we can also represent states as points

[13] This latter question is one that has been raised explicitly by Erik Winsberg and Arthur Fine, about whom more anon. See Winsberg and Fine [14].

[14] Note that $I\left(p^{A}|i,j\right)$ is a function not only of λ, but also of i and j. So this quantity cannot be interpreted as the information associated merely with λ. This "contextual" character of SEPARABILITY attaches to COMPLETENESS and CONDITION C as well. I leave as a project for future research the investigation of relationships that hold among suitable counterparts to these three constraints, where the counterparts are characterized in terms of λ alone.

in a phase space, connecting states with a set of exact values of the position and momentum of each particle in the system. Suppose we were to try to generalize this: consider theories for which the states are given by a specification of exact values of some distinguished set of physical magnitudes. We might even choose to call these quantities "elements of physical reality", or "EPRs".[15] These EPRs need not be restricted to properties like the values of position and momentum, nor need the EPRs be restricted to properties in the ordinary sense at all. They might include, for example, irreducibly dispositional quantities, represented probabilistically as physical propensities. How then might we think about characterizing our "separability" constraint? This leads us back to question (i), and suggests something of the following sort:

EPR-Separability

A theory will be said to be *EPR-SEPARABLE* (or to satisfy *EPR-SEPARABILITY*) just in case it satisfies each of the following constraints:

(1) The theory distinguishes a collection of physical magnitudes, $\{\sigma_1, \sigma_2, \ldots, \sigma_K\}$, as the "elements of physical reality" or "EPRs", exact values of which uniquely specify the state. For all $\lambda \in \Lambda$, $\lambda = \lambda(<\sigma_1, \sigma_2, \ldots, \sigma_K>)$.

(2) For each state $\lambda \in \Lambda$, there exist unique subsystem states, λ_A and λ_B, such that $\lambda_A = \lambda_A(<\sigma_1^A, \sigma_2^A, \ldots, \sigma_K^A>)$, $\lambda_B = \lambda_B(<\sigma_1^B, \sigma_2^B, \ldots, \sigma_K^B>)$, and for each $k = 1, 2, \ldots, K$, $\sigma_k = \sigma_k(<\sigma_k^A, \sigma_k^B>)$.

(3) There exists a function $p_\omega(z|k_1, k_2)$ such that $p_\lambda^A(x|i,j) = p_{\lambda_A}(x|i,j)$ and $p_\lambda^B(y|i,j) = p_{\lambda_B}(y|j,i)$.

(4) $\bar{I}(\lambda|i,j) = \bar{I}(\lambda_A|i,j) + \bar{I}(\lambda_B|j,i)$, where

$$\bar{I}(\lambda|i,j) \equiv I(p_\lambda^{AB}|i,j) \equiv \sum_{x,y} p_\lambda^{AB}(x,y|i,j) \ln p_\lambda^{AB}(x,y|i,j)$$

and $\bar{I}(\lambda_\Gamma|k_1,k_2) \equiv \sum_z p_{\lambda_\Gamma}(z|k_m,k_{m'}) \ln p_{\lambda_\Gamma}(z|k_m,k_{m'})$, where

$$m = \begin{cases} 1, \Gamma = A \\ 2, \Gamma = B \end{cases} \quad \text{and } m' = \begin{cases} 2, \Gamma = A \\ 1, \Gamma = B \end{cases}$$

(5) $p_\lambda^{AB}(x,y|i,j) = p_{\bar{\lambda}}^{AB}(y,x|j,i)$, where $\bar{\lambda} = \lambda(<\bar{\sigma}_1, \bar{\sigma}_2, \ldots, \bar{\sigma}_n>)$ and for each $k = 1, 2, \ldots, K$, $\bar{\sigma}_k = \sigma_k(<\sigma_k^B, \sigma_k^A>)$.

(6) $\rho(\lambda) = \phi_A(\lambda_A)\phi_B(\lambda_B)$, where $\phi_A = \phi_B \equiv \phi$.

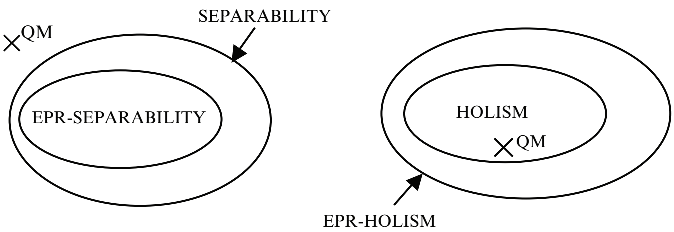
(7) $\Lambda = \Lambda_A \times \Lambda_B$, where $\Lambda_A = \Lambda_B \equiv \Omega$.

(8) $P^{AB}(x,y|i,j) \equiv \int_\Lambda \rho(\lambda) p_\lambda^{AB}(x,y|i,j) d\lambda$ and $P^\Gamma(z|k_1,k_2) \equiv \int_{\Lambda_\Gamma} \phi_\Gamma(\lambda_\Gamma) p_{\lambda_\Gamma}(z|k_m, k_{m'})\, d\lambda_\Gamma$, where m and m' are defined as in clause (4) above.

[15] Every ray of a Hilbert space, and so, every pure state of quantum mechanics, corresponds uniquely to some ("complete", in the quantum-mechanical sense) set of eigenvalues of a maximal set of commuting Hermitian operators. However, no one such set of Hermitian operators will serve for all pure states. So the states posited by quantum mechanics defy representation in terms of EPRs.

Clauses (1) and (2) of EPR-SEPARABILITY require that each state, for both the composite and the component systems, is fixed by appropriate sets of EPRs as indicated, so that each composite system state λ is decomposable into a pair of component system states, $\lambda_{A,}$ to be associated with the A subsystem, and λ_B, to be associated with the B subsystem. Clause (3) requires that the probability functions determined by the subsystem states λ_A and λ_B, respectively, be equal to the corresponding marginal probability functions, as is needed for consistency with the definition of the joint probability functions associated with the composite system states. It also requires that each EPR associated with λ_A "play the same role" as the corresponding EPR associated with λ_B in generating their respective probability functions. Clause (4) is my proposal for imposing on this picture what I consider to be a central feature of any acceptable separability constraint. It is the requirement that the Shannon information associated with the probability function determined by the composite system state description be equal to the sum of the two corresponding quantities determined by the subsystem states. *The whole is equal to the sum of the parts.* SEPARABILITY *simpliciter* makes the analogous demand without reference to any EPR-conception of physical states.[16]

From clause (4), EPR-SEPARABILITY entails plain SEPARABILITY, while from clauses (1) and (2), the converse does not hold. The point here is simply that a theory may satisfy SEPARABILITY without positing anything whatsoever regarding "elements of physical reality" that provide a decomposition of a λ into a λ A and a λ B. (See the figure below.) If one prefers to think about this from the other direction, we might define HOLISM simply as the denial of SEPARABILITY, and EPR-HOLISM as the denial of EPR-SEPARABILITY.



EPR–SEPARABILTY $\underset{\not\Leftarrow}{\overset{\Rightarrow}{}}$ SEPARABILITY

EPR–HOLISM $\equiv$ ~ (EPR–SEPARABILITY)

HOLISM $\equiv$ ~ SEPARABILITY

HOLISM $\underset{\not\Leftarrow}{\overset{\Rightarrow}{}}$ EPR–HOLISM

[16] Consequently, I presume that of those attracted to an information-theoretic characterization of "separability", those with a realist bent would favor something akin to EPR-SEPARABILITY, while those situated more toward the anti-realist end of the spectrum would prefer SEPARABILITY.

If we proceed, at least for the moment, under the assumption that we must reject SEPARABILITY, then this has consequences that are not as widely appreciated as I believe they ought to be. I will elaborate on this shortly. Returning to the characterization of EPR-SEPARABILITY, clauses (5), (6), and (7) impose some symmetry conditions on the two component systems based on the identity of the constituent entities, while clause (8) gives the standard prescription for generating testable predictions by summing over all states.[17]

If COMPLETENESS is to be abandoned, it should be of some interest to consider the relationship between this constraint and SEPARABILITY. As it happens, the two are logically equivalent.[18] This gives us the following chain of equivalences:[19]

SEPARABILITY

$$\Updownarrow$$

$(*)\forall\lambda, i, j,$

$$c_{11}{}^{c_{11}} c_{12}{}^{c_{12}} c_{21}{}^{c_{21}} c_{22}{}^{c_{22}} = (c_{11} - \Delta)^{c_{11}} (c_{12} + \Delta)^{c_{12}} (c_{21} + \Delta)^{c_{21}} (c_{22} - \Delta)^{c_{22}},$$
$$\text{where } \Delta \equiv \det C(\lambda|i,j) = c_{11}c_{22} - c_{12}c_{21}$$

$$\Updownarrow$$

$\forall\lambda, i, j,\ \Delta = 0$

$$\Updownarrow$$

CONDITION C

$$\Updownarrow$$

COMPLETENESS

Thus, in rejecting COMPLETENESS, we give up not just EPR-SEPARABILITY, but also that gap between EPR-SEPARABILITY and (the logically weaker) SEPARABILITY. Depending on one's views about the proper way to characterize states, one might say that the implications of Bell's Theorem and the experimental tests of the Bell Inequalities are even more drastic than previously thought. Our retreat

---

[17] For the sake of simplicity, I forego the readily available revisions needed to accommodate non-identical particles.

[18] This is an immediate consequence of the Gibbs Inequality (sometimes called "the Shannon-Gibbs Inequality". See Ash [15], pp. 16–19, for a proof. The Gibbs Inequality is a special case of a more general result that applies to all convex functions. For this more general result, see Billingsley [16], pp. 544–545. I am grateful to David Malament for calling my attention to the latter.

[19] I will have some things to say presently about the constraint denoted by the asterisk. For now, note that it may be seen to arise as follows: express SEPARABILITY in terms of the elements of the matrix $C(\lambda|i,j)$, and exponentiate both sides of the resulting equation. $(^*)$ emerges after a further bit of algebraic manipulation. That CONDITION C (and, so, COMPLETENESS) implies SEPARABILITY is now trivial. (The implication in the other direction is not.) For a generalization of each of these constraints, see Appendix 2.

from the classical worldview is even more extreme than perhaps is generally recognized. I wish to emphasize this point because the conventional wisdom has been that the deep meaning of Bell's Theorem and related experimental work is that not only deterministic theories, but even indeterministic theories that are not suitably "holistic", must be given up once and for all; and one *might* have thought that the holism in question, that associated with the "passion-at-a-distance" exhibited by the entangled states of systems in Bell-type experiments, in violation of COMPLETENESS, is really no more than that which follows from rejecting an EPR-style assumption of separability. In particular, one might have thought that every theory that satisfies COMPLETENESS (or, equivalently, SEPARABILITY), even those that do *not* posit states that are represented in terms of EPRs, is empirically equivalent to some such theory. If this were so, that logical gap between EPR-SEPARABILITY and SEPARABILITY *simpliciter* would be empirically vacuous, and most (certainly those who incline toward any form of anti-realism) would decline to acknowledge any meaningful distinction between the two classes of theories. However, it can be shown that there do exist COMPLETE (or, if one prefers, SEPARABLE) theories that are empirically equivalent to ***no*** EPR-SEPARABLE theory. (See Appendix 1.)

Consider that from a given composite-system state λ, COMPLETENESS alone determines neither λ_A nor λ_B. While it does afford a factorization of the joint probabilities into a product of the marginals, it remains utterly indifferent with regard to the manner of specification of the separate states themselves. Unless one simply rejects the EPR conception of state by *fiat* (in which case the marginal probabilities may be identified with the subsystem states), EPR-SEPARABILITY appears as a reasonable demand; and the constraint that states be represented in the form codified via that demand induces an *empirical* restriction that (merely) SEPARABLE theories need not meet.

I think it helpful to note that SEPARABILITY (and, so, COMPLETENESS) requires only that the ***quantity*** of information associated with the probability function determined by the composite system state description equal the sum of the corresponding ***quantities*** associated with the two marginal probability functions. But given this EPR conception of physical states, that requirement can well be met, even though the chunks of information appearing on the two sides of the equation that defines SEPARABILITY are ***qualitatively*** distinct.[20] Moreover, such qualitative distinctions can have empirical consequences. It is striking that empirical tests of the Bell Inequalities appear to demand that we reject even the weaker constraint (SEPARABILITY), the one having to do with information in this merely quantitative sense. This is what I meant when I suggested previously that Bell's Theorem and the associated experimental results force us to a more radical retreat from the classical world view than is perhaps generally recognized. I take EPR-SEPARABILITY to be a natural generalization of the classical worldview in the spirit of Einstein, Podolsky, and Rosen. It jettisons determinism as a requirement, but it retains the fundamental role of elements of physical reality. And yet, we see that the holism we

[20] For this insight I am indebted to Brandon Fogel, who suggested something similar to it in conversation.

are compelled to embrace is even more extreme than that which follows from the mere rejection of this EPR-SEPARABILITY.

A few years ago, Winsberg and Fine proposed that we take as a suitable "separability" constraint, that the joint probability function (associated with the composite system) simply be some function of the marginals.[21] That it be the *product* of the marginals would then be a simple special case of that proposal; but to insist that the function in question take that *particular* form would be overly restrictive, on their view. However, it can now be seen that if one agrees to adopt the information theoretic conception of "separability", and if one chooses to characterize states as probability functions (***whether or not*** those probability functions are taken to be the manifestations of some set of more fundamental elements of physical reality), then the "choice" of the product function is forced upon us.

One further point about the equivalence of SEPARABILITY and COMPLETENESS: We can define *mutual information* as indicated here:

Mutual Information

$$\forall \lambda, i, j,$$

$$\begin{aligned} I[X;Y]_{\lambda,i,j} &\equiv \sum_{\substack{x=\pm 1\\ y=\pm 1}} p_\lambda^{AB}(x,y|i,j) \ln\left[\frac{p_\lambda^{AB}(x,y|i,j)}{p_\lambda^{A}(x|i,j)\, p_\lambda^{B}(y|i,j)}\right] \equiv I[Y;X]_{\lambda,i,j} \\ &= \sum_{\substack{x=\pm 1\\ y=\pm 1}} p_\lambda^{AB}(x,y|i,j) \ln\left[\frac{p_\lambda^{A}(x|i,j,y)}{p_\lambda^{A}(x|i,j)}\right] \\ &= \sum_{\substack{x=\pm 1\\ y=\pm 1}} p_\lambda^{AB}(x,y|i,j) \ln\left[\frac{p_\lambda^{B}(y|i,j,x)}{p_\lambda^{B}(y|i,j)}\right], \end{aligned}$$

$$(\text{for } p_\lambda^{A}(x|i,j)\, p_\lambda^{B}(y|i,j) \neq 0)$$

NOTE: COMPLETENESS $\Rightarrow \forall \lambda, i, j, I[X;Y]_{\lambda,i,j} = I[Y;X]_{\lambda,i,j} = 0$

That mutual information is a symmetric function is apparent from the definition. The conditional probabilities appearing in the definition reduce to the unconditioned probabilities if and only if COMPLETENESS is satisfied, so COMPLETENESS holds just in case the mutual information is zero. That means, in turn, that (for a specified pair of detector switch settings) the information in the composite system state description is such that conditionalization on the outcome at one measurement site has no bearing on the probabilities assigned at the other site, the former merely providing redundant information with respect to such a "complete" state description.

---

[21] See Winsberg and Fine [14]. These authors challenge the views developed by Howard. For a trenchant analysis of these issues, see Fogel [17].

I note this in an effort to add reinforcement to the earlier interpretation I suggested for COMPLETENESS, where I also used the word "information", but there in an informal, non-technical sense.

Consider now the following constraint:[22]

Condition S

$$\forall \lambda, i, j,$$
$$c_{11}^{c_{11}} c_{12}^{c_{12}} c_{21}^{c_{21}} c_{22}^{c_{22}} = (c_{11} - \Delta)^{c_{11}} (c_{12} + \Delta)^{c_{12}} (c_{21} + \Delta)^{c_{21}} (c_{22} - \Delta)^{c_{22}},$$

where $\Delta \equiv \det C(\lambda | i, j)$.

Suppose an imaginary someone who has a reasonably good familiarity with the philosophical literature on Bell's Theorem, but who is blissfully unaware of any attempts toward an information-theoretic construal of "separability". What might be this person's initial response upon being told that COMPLETENESS and SEPARABILITY are logically equivalent? I daresay it might be something such as this, "So, I see that someone has succeeded in expressing COMPLETENESS in a grossly more complicated and obscure way." This might be followed by a sarcastic "(S)he has my congratulations." However, if our imaginary friend were to reflect for a time and play around with CONDITION S sufficiently, the response might continue in this manner, "Wait a minute, that… that looks like entropy, I mean that's a constraint on information; hey, that's a requirement that the information associated with the probability function representing the composite system is equal to the sum… Could we interpret that as a 'separability' condition of some sort?" Our friend just might discern this link between the two conceptually distinct classes of theories: the class of theories that satisfy COMPLETENESS on the one hand, and the class of theories that satisfy SEPARABILITY on the other. That those two sets are coextensive might be deemed to be of genuine significance.

This exercise of the imagination is intended to set the stage for a little speculation. Might something similar to that which our friend achieved with CONDITION S be done for CONDITION C as well? Of course, there is no reason why that would have to be so; CONDITION C might be nothing more than a way of expressing COMPLETENESS (or, equivalently, SEPARABILITY) in a different mathematical form; it might carry with it no suitably independent physical interpretation that would help us to enrich our understanding of COMPLETENESS and SEPARABILITY. However, the fact that CONDITION C is posed as the vanishing of a determinant is very suggestive, even tantalizing; for this allows us to recognize CONDITION C as the necessary and sufficient condition that there exist a non-trivial solution to a particular set of equations.

---

[22] The alert reader will immediately recognize CONDITION S as the aforementioned asterisked constraint.

CONDITION C
$$\Updownarrow$$
$\forall \lambda, i, j$, the following pair of equations has a non-trivial solution:

$$z_1 p_\lambda^{AB}(+1,+1|i,j) + z_2 p_\lambda^{AB}(+1,-1|i,j) = 0$$
$$z_1 p_\lambda^{AB}(-1,+1|i,j) + z_2 p_\lambda^{AB}(-1,-1|i,j) = 0$$

[equivalently, $\forall \lambda, i, j$, the following pair of equations has a non-trivial solution:

$$z_1 p_\lambda^{AB}(+1,+1|i,j) + z_2 p_\lambda^{AB}(-1,+1|i,j) = 0$$
$$z_1 p_\lambda^{AB}(+1,-1|i,j) + z_2 p_\lambda^{AB}(-1,-1|i,j) = 0$$

Alternatively,

$$\text{CONDITION C} \Longleftrightarrow \begin{cases} \forall \lambda, i, j, \text{ there exists a zero} - \text{eigenvalue} \\ \text{solution to the equation} \\ Cz = \gamma z \end{cases}$$

We can work out the solutions to these equations and put them in various forms and try to interpret them.[23] We can also note that CONDITION C asserts the existence of a zero-eigenvalue solution to a certain matrix equation. Does that matrix equation have an independent physical significance? What would it mean to demand of a theory that the corresponding matrix equation of that sort have *any* solution, let alone that it have a zero-eigenvalue solution?[24]

Success in the search for an independent physical interpretation of CONDITION C might yet deliver to us at least some small deepening of our understanding of COMPLETENESS and SEPARABILITY, and thereby help us to achieve a firmer grasp of precisely what it is that Bell's Theorem has to teach us about the way the world is not.

Acknowledgements I have had the great good fortune to discuss many of the ideas presented in this paper with generations of colleagues and students, my mentors and friends. It is no longer possible for me to give an accurate accounting of the many valuable contributions each has made to this work. However, I owe special thanks to Thomas Ferguson and Brandon Fogel, each of whom

[23] However these solutions are expressed, they suggest a symmetry condition of some sort (as does CONDITION C itself).

[24] For the quantum-mechanical case of two spin-1/2 particles in the singlet state, we have $c_{11} = c_{22} = (1/2)\sin^2(\theta/2)$, and $c_{12} = c_{21} = (1/2)\cos^2(\theta/2)$, where θ is the angle between the axes of alignment of the Stern-Gerlach magnets at A and B. This yields solutions to $Cz = \gamma z$ for eigenvalues $\gamma^{\pm} = (1/2)\sin^2(\theta/2)$, and $c_{12} = c_{21} = (1/2)\,[\sin^2(\theta/2) \pm \cos^2(\theta/2)]$. The eigenvectors associated with the γ^- eigenvalue have a curious similarity in form to those associated with the $\gamma^- = 0$ eigenvalue solutions to $Cz = \gamma z$ for the states posited by theories that satisfy CONDITION C (COMPLETENESS, SEPARABILITY). Perhaps this similarity in form will permit us to interpret CONDITION C as a symmetry constraint that quantum mechanics itself satisfies in some less restrictive, more general form. Perhaps.

has helped me to clarify my thinking in crucial ways. The value of the assistance rendered by Thomas Ferguson in working through all of the details and in preparing the manuscript has been inestimable.

Appendix 1

CLAIM: There exist SEPARABLE theories that are empirically equivalent to no EPR-SEPARABLE theory.

To establish the CLAIM, it will be helpful first to show the following:

$$(^{**})\ \text{EPR-SEPARABILITY} \Longrightarrow \begin{cases} \forall x, y, i, j, \\ P^{AB}(x,y|i,j) = P^{AB}(y,x|j,i) \end{cases}$$

Note the following, straightforward consequence of EPR SEPARABILITY:

$$\forall \lambda, x, y, z, i, j,$$

(1) $p_\lambda^{AB}(x,y|i,j) = p_\lambda^A(x|i,j)p_\lambda^B(y|i,j)$
(2) $\overline{\Lambda} = \Lambda$ and $\rho(\lambda) = \rho(\overline{\lambda})$
(3) $P^{AB}(x,y|i,j) = P^A(x|i,j)P^B(y|i,j)$
(4) $P^A(z|i,j) = P^B(z|j,i)$

For arbitrary values of x,y, i, and j,[25]

$$\begin{aligned}
P^{AB}(x,y|i,j) &= P^A(x|i,j)P^B(y|i,j) && (3)\\
&= \int_\Omega \phi(\omega)p_\omega(x|i,j)d\omega \int_\Omega \phi(\omega')p_{\omega'}(y|j,i)d\omega' && <7>,<8>\\
&= \int_{\Lambda_B} \phi_B(\lambda_B)p_{\lambda_B}(x|i,j)d\lambda_B \int_{\Lambda_A} \phi_A(\lambda_A)p_{\lambda_A}(y|j,i)d\lambda_A && <7>\\
&= \int_{\Lambda_B\times\Lambda_A} \phi_B(\lambda_B)p_{\lambda_B}(x|i,j)\phi_A(\lambda_A)p_{\lambda_A}(y|j,i)d\lambda_B d\lambda_A\\
&= \int_{\Lambda_B\times\Lambda_A} \phi_B(\lambda_B)\phi_A(\lambda_A)p_{\lambda_B}(x|i,j)p_{\lambda_A}(x|i,j)d\lambda_B d\lambda_A\\
&= \int_{\bar{\Lambda}} \rho(\overline{\lambda})p_{\overline{\lambda}}^A(x|i,j)p_{\overline{\lambda}}^B(y|j,i)d\overline{\lambda} && (2),<3>,<6>
\end{aligned}$$

[25] In what follows, numbers in angle brackets refer to the corresponding clauses in the definition of EPR-SEPARABILITY, while numbers in parentheses refer to the corresponding clauses of the aforementioned consequence of EPR-SEPARABILITY.

$$= \int_{\bar{\Lambda}} \rho(\overline{\lambda}) p^{AB}_{\overline{\lambda}}(x,y|i,j) d\overline{\lambda} \quad (1)$$

$$= \int_{\Lambda} \rho(\lambda) p^{AB}_{\overline{\lambda}}(x,y|i,j) d\lambda \quad (2)$$

$$= \int_{\Lambda} \rho(\lambda) p^{AB}_{\lambda}(y,x|j,i) d\lambda \quad <5>$$

$$= P^{AB}(y,x|j,i) \quad <8>$$

This establishes (**).

Now let T be a theory that posits a set of states $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ and uniform distribution function $\rho(\lambda_i) = \frac{1}{m}$ and suppose that for each $i = 1, 2, \ldots, m$, there are values α_i, β_i, γ_i such that for all k:

$$C(\lambda_i|k,k) = \frac{1}{(1+\gamma_i)(\alpha_i+\beta_i)} \begin{bmatrix} \alpha_i & \beta_i \\ \gamma_i\alpha_i & \gamma_i\beta_i \end{bmatrix}, \text{ where } \beta_i > \gamma_i\alpha_i.$$

(Note: These states are normalized, and T satisfies CONDITION C.)

For T, then,

$$\forall k, P^{AB}(1,-1|k,k) = \frac{1}{m}\sum_{i=1}^{m} \frac{\beta_i}{(1+\gamma_i)(\alpha_i+\beta_i)}, \text{ and}$$

$$P^{AB}(-1,1|k,k) = \frac{1}{m}\sum_{i=1}^{m} \frac{\gamma_i\alpha_i}{(1+\gamma_i)(\alpha_i+\beta_i)}$$

$$< \frac{1}{m}\sum_{i=1}^{m} \frac{\beta_i}{(1+\gamma_i)(\alpha_i+\beta_i)} = P^{AB}(1,-1|k,k)$$

Thus, by (**), T violates EPR-SEPARABILITY; but since T satisfies CONDITION C, it also satisfies SEPARABILITY. Moreover, since T requires that for all k, $P^{AB}(1,-1|k,k) \neq P^{AB}(-1,1|k,k)$, this SEPARABLE theory makes predictions that are compatible with no EPR-SEPARABLE theory.

This establishes the CLAIM.

Appendix 2

COMPLETENESS, SEPARABILITY, CONDITION C, and equation (*) all admit of generalization to cases with arbitrary values of $s \geq 2$ and $n \geq 2$, where s is the number of component systems and n is the number of possible outcomes at each detector site. In the general case, we have the following:

OUTCOME VARIABLES: $x_1, x_2, \ldots, x_s$
POSSIBLE VALUES FOR VARIABLE x_k: $x_{k1}, x_{k2}, \ldots, x_{kn}$
DETECTOR SETTING VARIABLE FOR DETECTOR AT A_k: i_k

JOINT PROBABILITY FUNCTIONS:

$$p_\lambda^{A_1 A_2 \ldots A_s}(x_1, x_2, \ldots, x_s | i_1, i_2, \ldots, i_s)$$

MARGINALS:

$$p_\lambda^{A_K}(x_k | i_1, i_2, \ldots, i_s) \equiv \sum_{\substack{x_j \in \mathrm{X}_j \\ j \neq k}} p_\lambda^{A_1 A_2 \ldots A_s}(x_1, x_2, \ldots, x_s | i_1, i_2, \ldots, i_s),$$

where $\mathrm{X}_j \equiv \{x_{j1}, x_{j2}, \ldots, x_{js}\}$

and $\forall j, m, m', m < m' \Rightarrow x_{jm} \leq x_{jm'}$

COMPLETENESS

$\forall \lambda, x_j, i_j,$

$$p_\lambda^{A_1 A_2 \ldots A_s}(x_1, x_2, \ldots, x_s | i_1, i_2, \ldots, i_s) = \prod_{k=1}^{s} p_\lambda^{A_K}(x_k | i_1, i_2, \ldots, i_s)$$

SEPARABILITY

$\forall \lambda, i_j,$

$$I\left(p_\lambda^{A_1 A_2 \ldots A_s} | i_1, i_2, \ldots, i_s\right) = \sum_{k=1}^{s} I(p_\lambda^{A_k} | i_1, i_2, \ldots, i_s)$$

CONDITION C

For all $1 \leq j < m \leq s$, and for all$1 \leq j\prime < m\prime \leq s$,

$$\det \begin{bmatrix} c_{k_1 k_2 \ldots k_j \ldots k_m \ldots k_s} & c_{k_1 k_2 \ldots k_j \ldots k_{m'} \ldots k_s} \\ c_{k_1 k_2 \ldots k_{j'} \ldots k_m \ldots k_s} & c_{k_1 k_2 \ldots k_{j'} \ldots k_{m'} \ldots k_s} \end{bmatrix} = 0,$$

where $c_{k_1 k_2 \ldots k_s}(\lambda | i_1, i_2, \ldots, i_s) = p_\lambda^{A_1 A_2 \ldots A_s}(x_{1(k1)}, x_{2(k2)}, \ldots, x_{s(ks)} | i_1, i_2, \ldots, i_s).$

CONDITION Q

$\forall \lambda, i_j, k_j,$

$$q_{k_1 k_2 \ldots k_s}(\lambda | i_1, i_2, \ldots, i_s) = 0,$$

where $q_{k_1k_2\ldots k_s}(\lambda|i_1,i_2,\ldots,i_s) =$

$$-\left\{c_{k_1k_2\ldots k_s}\sum_{j=1}^{s-1}\left[\frac{(s-1)!(-1)^j}{j!(s-j-1)!}\left(\sum_{<k'_1,k'_2,\ldots,k'_s>\in M(k_1,k_2,\ldots,k_s)} c_{k'_1k'_2\ldots k'_s}\right)^j\right]\right.$$

$$\left.+\sum_{m=1}^{s}\left[(c_{k_1k_2\ldots k_s})^{s-m}\sum_{j_m>j_{m-1}>\ldots>j_1}\prod_{i=1}^{m}\left(\sum_{<k'_1,k'_2,\ldots,k'_s>\in K_{j_i}}^{c_{k'_1k'_2\ldots k'_s}}\right)\right]\right\},$$

$$\begin{aligned} M(k_1,k_2,\ldots,k_s) &\equiv \{<k'_1,k'_2,\ldots,k'_s> \mid\ <k'_1,k'_2,\ldots,k'_s> \neq <k_1,k_2,\ldots,k_s>\},\\ K_j(k_1,k_2,\ldots,k_s) &\equiv \{<k'_1,k'_2,\ldots,k'_s> \mid k'_j = k_j \& <k'_1,k'_2,\ldots,k'_s> \\ &\qquad \neq <k_1,k_2,\ldots,k_s>\}, \end{aligned}$$

and for $m = 1$, $\displaystyle\sum_{j_m>j_{m-1}>\ldots>j_1}\prod_{i=1}^{m}\left(\sum_{<k'_1,k'_2,\ldots,k'_s>\in K_{j_i}}^{c_{k'_1k'_2\ldots k'_s}}\right) \equiv \sum_{j=1}^{s}\left(\sum_{<k'_1,k'_2,\ldots,k'_s>\in K_j}^{c_{k'_1k'_2\ldots k'_s}}\right).$

NOTE: The generalized form of $(^*)$ is as follows:

$$\prod_{k_1,k_2,\ldots,k_s} c_{k_1k_2\ldots k_s}^{c_{k_1k_2\ldots k_s}} = \prod_{k_1,k_2,\ldots,k_s}\left(c_{k_1k_2\ldots k_s} - q_{k_1k_2\ldots k_s}\right)^{c_{k_1k_2\ldots k_s}}$$

For $n = s = 2$, we have $q_{11} = q_{22} = \det C = -q_{12} = -q_{21}$, and CONDITION Q reduces to $(^*)$ in this case.

I offer the following assertion here without proof:

$\forall n \geq 2$, $\forall s \geq 2$, the following are logically equivalent:

COMPLETENESS
CONDITION C
SEPARABILITY

References

1. Mermin, N. David (1981) "Quantum mysteries for anyone." *Journal of Philosophy* **78**: 397–408.
2. Jarrett, Jon P. (1989) "Bell's Theorem: a guide to the implications." In [5] 60–79.
3. Einstein, Albert, Boris Podolsky, and Nathan Rosen (1935) "Can quantum-mechanical description of physical reality be considered complete?" *Physical Review* **47**: 777–780.
4. Bohr, Niels (1935) "Can quantum-mechanical description of physical reality be considered complete?" *Physical Review* **48**: 696–702.
5. Cushing, James T. and Ernan McMullin, editors (1989) *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem*. Notre Dame, IN: University of Notre Dame Press.
6. Fine, Arthur (1989) "Do correlations need to be explained?" In [5], 175–194.

7. van Fraassen, Bas (1989) “The Charybdis of realism: epistemological implications of Bell’s Inequality.” In [5], 97–113.
8. Hughes, R. I. G. (1989) “Bell’s Theorem, ideology, and structural explanation.” In [5], 195–207.
9. Fine, Arthur (1986) *The Shaky Game: Einstein, Realism, and the Quantum Theory*. Chicago: University of Chicago Press.
10. Einstein, Albert (1936) “Physik und Realität.” *Journal of the Franklin Institute* **221**: 313–347.
11. Einstein, Albert (1948) “Quantenmechanik und Wirklichkeit.” *Dialectica* **2**: 320–324.
12. Howard, Don (1985) “Einstein on locality and separability.” *Studies in History and Philosophy of Science* **16**: 171–201.
13. Howard, Don (1989) “Holism, separability, and the metaphysical implications of the Bell experiments.” In [5], 224–253.
14. Winsberg, Eric, and Arthur Fine (2003) “Quantum life: interaction, entanglement, and separation.” *Journal of Philosophy*, **C(2)**: 80–97.
15. Ash, Robert B. (1990) *Information Theory*. New York: Dover Publications. (Originally published in 1965 by Interscience Publishers, New York.)
16. Billingsley, Patrick (1995) *Probability and Measure*, 3rd edition. New York: Wiley. (First edition published in 1979.)
17. Fogel, Brandon (2007) “Formalizing the separability condition in Bell’s Theorem”. *Studies in History and Philosophy of Modern Physics*, **38**, 920–937.