

Chapter 12

Quantum Mechanics in a Nutshell

12.1 The Quantum Canon

At the moment, there exists a loosely bundled canon of quantum rules subsumed under the term *quantum mechanics* or *quantum theory*. It includes reversible as well as irreversible processes, and is *prima facie* inconsistent. As already von Neumann [552, 554] and later Everett [30, 206, 545] noted, there cannot be any irreversible measurement process nested in a ubiquitous uniformly reversible evolution of the quantum state. Both von Neumann and Everett called the former, irreversible, discontinuous change the “*process 1*”; and the latter, reversible, continuous, deterministic change the “*process 2*,” respectively. Stated differently, there cannot exist any irreversible many-to-one measurement scenario (other than pragmatic fappness) in a reversible one-to-one environment.

Hence, if one wants to maintain irreversible measurements, then (at least within the quantum formalism) one is faced with the following dilemma: either quantum mechanics must be augmented with some irreversible, many-to-one state evolution, thereby spoiling the ubiquitous, universal reversible one-to-one state evolution; or the assumption of the co-existence of a ubiquitous, uniform reversible one-to-one state evolution on the one hand with some irreversible many-to-one “wave function collapse,” (by another wording, “reduction of the state vector”) throughout measurement on the other hand, yields a complete contradiction.

How is such a situation handled in other areas? Every system of logic which is self-contradictory (inconsistent) – such that a proposition as well as its negation is postulated; or can be derived from the postulates – in particular, in a formal axiomatic system, is detrimental and disastrous. Because by the *principle of explosion* (Latin: *ex falso quodlibet*) any invocation of a statement as well as of its negation yields every proposition true. This can be motivated by supposing that both “*P*” as well as “not *P*” are true. Then the proposition “*P* or anything” is true (because at least “*P*” is true). Now suppose that also “not *P*” holds. But then, in order for “*P* or anything” to be true, “anything” needs to be true. However, if anything is derivable, then such a system lacks any descriptive or predictive capacity. In this

respect it is quite convenient that quantum mechanics does not represent a formal system in the strict logical sense.

With regards to the persistence and scientific reception of inconsistencies within theoretical domains one is reminded of Cantorian “naive” set theory [116, 117]; whereby a set, or aggregate, was defined as follows [118, p. 85]: “By an “aggregate” (Menge) we are to understand any collection into a whole (Zusammenfassung zu einem Ganzen) M of definite and separate objects m of our intuition or our thought. These objects are called the “elements” of M .” Despite its well known inconsistencies (e.g., Russell’s paradox, [288] defining a “set of all sets that are not members of themselves”), it was embraced by researchers of the time with unabated enthusiasm. Hilbert, for instance, stated that [278] “Wherever there is any hope of salvage, we will carefully investigate fruitful definitions and deductive methods. We will nurse them, strengthen them, and make them useful. No one shall drive us out of the paradise which Cantor has created for us.” Indeed, the different forms of (un)countable infinities still present a marvel of early “naive” set theory.

Another source of perplexity remains irreversibility in statistical physics [381]; in particular, issues related to the second law of thermodynamics [375] in view of microphysical irreversibility. As already pointed out in Sect. 1.1, for the second law of thermodynamics to hold Maxwell advised to avoid [234, p. 422]: “all personal enquiries [[of Molecules]] which would only get me into trouble.” A recent discussion [84, 158, 380, 431] on the exorcism of Maxwell’s demon [189, 190, 332] is witness of the ongoing debate.

Many practitioners either tend to look the other way, or take a pragmatic stance expressed quite voluptuously by Heaviside [272, Sect. 225]: “I suppose all workers in mathematical physics have noticed how the mathematics seems made for the physics, the latter suggesting the former, and that practical ways of working arise naturally. . . . But then the rigorous logic of the matter is not plain! Well, what of that? Shall I refuse my dinner because I do not fully understand the process of digestion? No, not if I am satisfied with the result. Now a physicist may in like manner employ unrigorous processes with satisfaction and usefulness if he, by the application of tests, satisfies himself of the accuracy of his results. At the same time he may be fully aware of his want of infallibility, and that his investigations are largely of an experimental character, and may be repellent to unsympathetically constituted mathematicians accustomed to a different kind of work.”

12.2 Assumptions of Quantum Mechanics

As suggested by Dirac [173] and explored by von Neumann [552, 554], quantum mechanics has been formalized in terms of Hilbert spaces.

Many researchers have attempted to at least partially derive this kind of quantum formalism from other principles, mostly informational (cf., e.g., Refs. [569, 588], and [239, Part II], to name but a few). Indeed, as Lakatos has pointed out [324], the contemporary researchers cannot know which ideas will prevail, and will ultimately

result in progressive research programs. Therefore it appears prudent to pursue varied research programs in parallel.

In the following we shall present a very brief, somewhat revisionist, view on quantum mechanics. It is based on pure *quantum states* representable as dichotomic value assignments on, equivalently, a (normalized) system of orthonormal basis vectors, the associated set of projection operators, or the associated set of subspaces of a Hilbert space. (Fapp a Hilbert space is a vector space with a scalar product.) Vector spaces are needed for the manipulation of vectors, such as vector additions and superpositions. (For the rest of this chapter, suppose that we are “riding” a single vector of a high dimensional Hilbert space, thereby qualifying as “members of the church of the larger Hilbert space.”)

By the spectral theorem, *observables* can be represented by the weighted spectral sums of such pure (mutually orthogonal) quantum states as well. Any non-degenerate spectral sum represents a maximal measurement. We may call this, or rather the set of orthogonal projection operators in the spectral sum, a *context*.

Quantum *complementarity* is the feature that two different contexts cannot be directly measured simultaneously.

Scalar products are needed for defining the relational property of vectors, such as orthogonality and collinearity. They allow projections of vectors onto arbitrary non-zero subspaces. Thereby they grant a particular view on the quantum state, as seen from another quantum state – or, equivalently, the proposition represented by the respective vector or associated projection operator.

Ultimately, scalar products facilitate the definition of frame functions which can be interpreted as *quantum probabilities*. This is necessary because, at least from dimension three onwards, the tight intertwining (pasting) of such maximal views or contexts does not allow quantum probabilities to be defined by the convex sum of two-valued measures. These two-valued measures could, if they existed, be interpreted as non-contextual truth assignments. As it turns out, relative to reasonable side assumptions, any such classical strategy fails, simply because, from dimension three onwards, such two-valued measures do not exist for more than a *single* context.

12.3 Representation of States

Suppose we are given a Hilbert space of sufficient dimension. That is, its dimension coincides with the maximal number of mutually exclusive outcomes of any experiment we wish to formalize.

It is “reasonable” to define a physical state of an object by the *maximal empirical (information) content* in principle accessible to an observer by any sort of operational means available to this observer. In Dirac’s words [173, pp. 11–12], “A state of a system may be defined as an undisturbed motion that is restricted by as many conditions or data as are theoretically possible without mutual interference or contradiction. In practice the conditions could be imposed by a suitable preparation of the system,

consisting perhaps in passing it through various kinds of sorting apparatus, such as slits and polarimeters, the system being left undisturbed after the preparation.”

Schrödinger, in his *Generalbeichte* [452, Footnote 1, p. 845] (general confession) of 1935, pointed out that [539, Sect. 6, p. 328] “*Actually [[in truth]]—so they say—there is intrinsically only awareness, observation, measurement. If through them I have procured at a given moment the best knowledge of the state of the physical object that is possibly attainable in accord with natural laws, then I can turn aside as meaningless any further questioning about the “actual state,” inasmuch as I am convinced that no further observation can extend my knowledge of it—at least, not without an equivalent diminution in some other respect (namely by changing the state, see below).*”¹

No further justification is given here.

A *quantum state* is thus identified with a maximal co-measurable (or co-preparable) entity. This is based on *complementarity*: not all conceivable quantum physical properties are co-measurable. (For classical models of complementarity, see, for instance, Moore’s discrete-valued automaton analogue of the Heisenberg uncertainty principle [373, 446, 499], as well as Wright’s generalized urn model [578], and partition logics in general [511].)

In the Hilbert space formulation of quantum mechanics a state is thus formalized by two entities; some structural elements, and a measure on these elements [520]:

- (I) equivalently,
- (i) an orthonormal basis of Hilbert space;
 - (ii) a set of mutually orthogonal projection operators corresponding to an orthonormal basis called *context*;
 - (iii) a maximal observable, or maximal operator, or maximal transformation whose spectral sum contains the set of mutual orthogonal projection operators from the aforementioned basis;
 - (iv) a maximal Boolean subalgebra [249, 300, 376, 420] of the quantum logic also called a *block*;
- (II) as well as a two-valued (0-1) measure (or, used synonymously, *valuation*, or *truth assignment*) on all the aforementioned entities, singling out or selecting one of them such that this measure is one on exactly one of them, and zero on all the others.

¹German original [452, Sect. 6, p. 823] “*Wirklich – so sagt man – sind ja eigentlich nur Wahrnehmung, Beobachtung, Messung. Habe ich mir durch sie in einem gegebenen Augenblick die bestmögliche Kenntnis vom Zustande des physikalischen Objekts verschafft, die naturgesetzlich erlangbar ist, so darf ich jede darüber hinausgehende Frage nach dem “wirklichen Zustand” als gegenstandslos abweisen, sofern ich überzeugt bin, dass keine weitere Beobachtung meine Kenntnis davon erweitern kann – wenigstens nicht, ohne sie in anderer Hinsicht um ebensoviel zu schmälern (nämlich durch Veränderung des Zustandes, s. w. u.)*”.

Another way of formalizing a state would be to single out a particular vector of the basis referred to earlier – the one which is actually “true;” that is, whose measurement (deterministically) indicates that the system is in this state.

However, one cannot “not measure” the accompanying context of a particular set of orthogonal vectors which, together with the state vector, completes a basis. One can deny it, or look the other way, but the permutation quantum evolution presented below presents no way for “blissful ignorance:” any “beam dump” is fapp irreversible and only fapp formalizable by taking partial traces, whereas in principle the information about the rest of the context remains intact.

12.4 Representation of Observables

A non-degenerate *quantum observable* is identified with all properties of a state, less the two-valued measure, and formalized by

- (i) an orthonormal basis of Hilbert space;
- (ii) a set of mutually orthogonal projection operators corresponding to an orthonormal basis called *context*;
- (iii) a maximal observable, or maximal operator, or maximal transformation whose spectral sum contains the set of mutual orthogonal projection operators from the aforementioned basis;
- (iv) a maximal Boolean subalgebra [249, 300, 376, 420] of the quantum logic also called *block*.

This correspondence (ex measure) between a quantum state and a quantum observable is reflected in the formalism itself: Any maximal observable can be decomposed into a spectral sum, with the orthogonal projection operators forming a corresponding orthonormal basis, or, synonymously, by a context or a block.

12.5 Dynamical Laws by Isometric State Permutations

The *isometric state permutation rule* postulates that the quantum state evolves in a deterministic way by isometric (length preserving) state permutation. [Throughout this book we shall denote a bijection between the same set (continuum) as permutation.] This can be equivalently understood as a linear transformation preserving the inner product, or as change of orthonormal bases/contexts/blocks [260, Sect. 74] (see also [460]). The formalization is in terms of unitary operators.

Suppose that the quantum mechanical (unitary) permutation is ubiquitous and thus valid universally. Then, stated pointedly, “reversibility rules.”

This assumption is strongly supported by a nesting argument [30, 31] first put forward by Everett, and later by Wigner [571] (cf. Sect. 1.7 on p. 10). Because it is

quite reasonable that any observing agent, when combined with the object this agent observes (including the cut/interface), should form a system that is quantized again; thereby implying a time evolution which is governed by isometric state permutation.

12.6 Disallowed Irreversible Processes

With the assumption of uniform validity of state the quantum evolution by isometric permutativity, many-to-one processes are excluded. In particular, formation of mixed states from pure states, as well as irreversible measurements, and the associated “state reduction” (or “wave function collapse”) contradict the isometric state permutation rule, and cannot take place in this regime.

12.6.1 Disallowed State Reduction

Usually a “state reduction” occurs during an irreversible measurement. It is associated with a transition from a state which is in a non-trivial coherent (or, by an equivalent term, linear) superposition – that is, a linear combination – a multiplicity of more than one states $\sum_{i=1}^{n>1} \alpha_i |i\rangle$ with normalization $\sum_{i=1}^{n>1} |\alpha_i|^2 = \sum_{i=1}^{n>1} \bar{\alpha}_i \alpha_i = 1$ into a *single* state $|k\rangle$, $1 \leq k \leq n$ with probability $|\alpha_k|^2 = \bar{\alpha}_k \alpha_k$. No one-to-one process such as a permutation can produce this n -to-1 transition.

12.6.2 Disallowed Partial Traces

Again any “generation” of a mixed state from pure states by “tracing out” certain components of the state is disallowed, since this amounts to a loss of information, and does not correspond to any invertible (reversible) transformation. Conversely, one could “purify” any mixed state, but this process is nonunique.

12.7 Superposition of States – Quantum Parallelism

Already Dirac referred to the *principle of superposition of states* [173, pp. 11–12], “*whenever the system is definitely in one state we can consider it as being partly in each of two or more other states. The original state must be regarded as the result of a kind of superposition of the two or more new states, in a way that cannot be conceived on classical ideas.*”

The superposition principle can be formalized by linear combinations as follows: suppose two states, which can be formally represented by orthonormal bases

$\mathfrak{B} = \{|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle, \dots, |\mathbf{e}_n\rangle\}$ and $\mathfrak{B}' = \{|\mathbf{f}_1\rangle, |\mathbf{f}_2\rangle, \dots, |\mathbf{f}_n\rangle\}$. Then each member $|\mathbf{e}_i\rangle$ of the first basis can be represented as a *linear combination* or *coherent superposition* or *superposition* of elements of the second basis by

$$|\mathbf{e}_i\rangle = \sum_{j=1}^n \alpha_{ij} |\mathbf{f}_j\rangle; \quad (12.1)$$

and *vice versa*.

For normalization reasons which are motivated by probability interpretations, the absolute squares of the coefficients α_{ij} must add up to 1; that is,

$$\sum_{j=1}^n |\alpha_{ij}|^2 = \sum_{j=1}^n \overline{\alpha_{ij}} \alpha_{ij} = 1. \quad (12.2)$$

With this normalization, the dyadic (tensor) product of $|\mathbf{e}_i\rangle$ is always of trace class one; that is,

$$\begin{aligned} \text{Tr}(|\mathbf{e}_i\rangle\langle\mathbf{e}_i|) &= \text{Tr} \left[\left(\sum_{j=1}^n \alpha_{ij} |\mathbf{f}_j\rangle \right) \left(\sum_{k=1}^n \overline{\alpha_{ik}} \langle\mathbf{f}_k| \right) \right] = \\ &= \sum_{l=1}^n \langle\mathbf{f}_l| \left(\sum_{j=1}^n \alpha_{ij} |\mathbf{f}_j\rangle \right) \left(\sum_{k=1}^n \overline{\alpha_{ik}} \langle\mathbf{f}_k| \right) |\mathbf{f}_l\rangle = \\ &= \sum_{l,j,k=1}^n \alpha_{ij} \overline{\alpha_{ik}} \delta_{lj} \delta_{kl} = \sum_{l=1}^n \alpha_{il} \overline{\alpha_{il}} = 1. \end{aligned} \quad (12.3)$$

Superpositions of pure states – resulting in a pure state – should not be confused with mixed states, such as, for instance,

$$\rho = \sum_{j=1}^n \rho_{ij} |\mathbf{f}_i\rangle\langle\mathbf{f}_j|, \quad (12.4)$$

which are the linear combination of dyadic (tensor) products $|\mathbf{f}_i\rangle\langle\mathbf{f}_j|$ of pure states $|\mathbf{f}_i\rangle$ and $|\mathbf{f}_j\rangle$ such that $\text{Tr}(\rho) = 1$ and $\text{Tr}(\rho^2) < 1$.

12.8 Composition Rules and Entanglement

In classical physics any compound system – the whole – can be composed from its parts by separation and specification of the parts individually.

This “factoring” of states of multiple constituent parts into products of individual single particle states need no longer be possible in quantum mechanics (although it is not excluded in particular quasi-classical cases): in general, any strategy to obtain the entire state of the whole system of many particles by considering the states of the individual particles fails.

This is a consequence of the quantum mechanical possibility to *superpose states of multiple particles*; that is, to add together arbitrarily weighted (subject to normalization) products of single particle states to form a new, valid, state. Classically, these states are “unreachable” by reversible evolutions-by-permutation, but quantum mechanically it is quite straightforward to create such a superposition through unitary transformations. Arguably the most prominent one is a Hadamard transformation corresponding to a 50:50 beam splitter.

12.8.1 *Relation Properties About Versus Individual Properties of Parts*

Probably the first to discuss this quantum feature (in the context of the measurement process) was von Neumann, stating that, “If I is in the state $\varphi(q)$ and II in the state $\xi(r)$, then $I + II$ is in the state $\Phi(q, r) = \varphi(q)\xi(r)$. If on the other hand $I + II$ is in a state $\Phi(q, r)$ which is not a product $\varphi(q)\xi(r)$, then I and II are mixtures and not states, but Φ establishes a one-to-one correspondence between the possible values of certain quantities in I and in II . [554, Sect. VI.2, pp. 436–437] ... all “probability dependencies” which may exist between the two systems disappear as the information is reduced to the sole knowledge of ... the separated systems I and II . But if one knows the state of I precisely, as also that of II , “probability questions” do not arise, and then $I + II$, too, is precisely known [554, Sect. VI.2, p. 426].”² Unfortunately the translation uses the two English phrases “probability dependencies” as well as “probability questions” for von Neumann’s German expression “Wahrscheinlichkeitsabhängigkeit.” Maybe it would be better to translate these by “probabilistic correlations.”

In a series of German [452] and English [453, 455] papers Schrödinger emphasized that [539, Sect. 10, p. 332] “The whole is in a definite state, the parts taken individually are not.”³

²German original: “Ist I im Zustande $\varphi(q)$ und II im Zustande $\xi(r)$, so ist $I + II$ im Zustande $\Phi(q, r) = \varphi(q)\xi(r)$. Ist dagegen $I + II$ in einem Zustande $\Phi(q, r)$, der kein Produkt $\varphi(q)\xi(r)$ ist, so sind I und II Gemische, aber Φ stiftet eine ein-eindeutige Zuordnung zwischen den möglichen Werten gewisser Größen in I und in II . [554, Sect. VI.2, p. 232] ... bei alleiniger Kenntnis ... der getrennten Systeme I und II , gehen alle “Wahrscheinlichkeitsabhängigkeiten”, die zwischen denn beiden Systemen noch bestehen können, verloren. Wenn man aber sowohl den Zustand von I als auch denjenigen von II genau kennt, kommen “Wahrscheinlichkeitsabhängigkeiten” nicht in Frage, und man kennt auch $I + II$ genau [554, Sect. VI.2, p. 227].”

³German original [452, Sect. 10, p. 827] “Das Ganze ist in einem bestimmten Zustand, die Teile für sich genommen nicht.”

Both von Neumann and Schrödinger thought of this as a sort of a zero-sum game, very much like complementary observables: due to the scarcity and fixed amount of information which merely gets permuted during state evolution, one can either have total knowledge of the individual parts; with zero relational knowledge of the correlations and relations among the parts; or conversely one can have total knowledge of the correlation and relations among the parts; but know nothing about the properties of the individual parts. Stated differently, any kind of mixture between the two extremes can be realized for an ensemble of multiple particles or parts:

- (i) either the properties of the individual parts are totally determined; in this case the relations and correlations among the parts remain indeterminate,
- (ii) or the relations and correlations among the parts are totally determined; but then the properties of the individual parts remain indeterminate.

For classical particles only the first case can be realized. The latter case is a genuine quantum mechanical feature.

Everett expressed this by saying that, in general (that is, with the exception of quasi-classical states) [206], “*a constituent subsystem cannot be said to be in any single well-defined state, independently of the remainder of the composite system.*” The entire state of multiple quanta can be expressed completely in terms of *correlations* or *joint probability distributions* [365, 576], or, by another term, *relational properties* [587, 588], among observables belonging to the subsystems. As pointedly stated by Bennett [287] in quantum physics the possibility exists “*that you have a complete knowledge of the whole without knowing the state of any one part. That a thing can be in a definite state, even though its parts were not. . . . It’s not a complicated idea but it’s an idea that nobody would ever think of.*”

Schrödinger called such states in German *verschränkt*, and in English *entangled*. In the context of multiple particles the formal criterion for entanglement is that an entangled state of multiple particles (an entangled multipartite state) cannot be represented as a product of states of single particles.

12.8.2 “Breathing” In and Out of Entanglement and Individuality

The sort of “zero-sum game” mentioned earlier is complementary with regards to encoding information into relations-correlations *versus* individual properties: due to the scarcity and fixed amount of information which merely gets permuted during state evolution, one can either have total knowledge of the individual parts; with zero relational knowledge of the correlations and relations among the parts; or conversely one can have total knowledge of the correlation and relations among the parts; but then one learns nothing about the properties of the individual parts. Stated differently, any kind of mixture between the following two extremes can be realized for an ensemble of multiple particles or parts:

- (i) individuality: either the properties of the individual parts are totally determined; in this case the relations and correlations among the parts remain indeterminate; in probability theory one may say that the parts are *independent* [261, Sect. 45]
- (ii) entanglement: or the relations and correlations among the parts are totally determined; but then the properties of the individual parts remain indeterminate.

For classical particles only the first, individual, case can be realized. The latter, entangled, case is a genuine quantum mechanical feature.

Thereby, interaction entangles any formerly individual parts at the price of losing their individuality, and measurements on individual parts destroys entanglement and “enforces value-definiteness” of the individual constituent parts. Suppose one starts out with a factorable case. Then an entangled state is obtained by a unitary transformation of the factorable state. Its inverse transformation leads back from the entangled state to the factorable state; through a continuum of non-maximal entangled intermedium states. This may go back and forth – from individuality to entanglement and then back to individuality – an arbitrary number of times.

In purely formal terms; that is, on the syntactic level, this can be quite well understood: a pure state of, say, k particles with n states per particle can, be written as

$$\sum_{i_1, \dots, i_k=1}^n \alpha_{i_1, \dots, i_k} |\psi_{1, i_1}\rangle \dots |\psi_{k, i_k}\rangle = \sum_{i_1, \dots, i_k=1}^n \alpha_{i_1, \dots, i_k} |\psi_{1, i_1} \dots \psi_{k, i_k}\rangle, \quad (12.5)$$

and not

$$\sum_{i_1, \dots, i_k=1}^n a_{1, i_1} \dots a_{k, i_k} |\psi_{1, i_1}\rangle \dots |\psi_{k, i_k}\rangle = \sum_{i_1, \dots, i_k=1}^n a_{1, i_1} \dots a_{k, i_k} |\psi_{1, i_1} \dots \psi_{k, i_k}\rangle. \quad (12.6)$$

In particular, this is only valid if $\alpha_{i_1, \dots, i_k} = a_{1, i_1} \dots a_{k, i_k}$.

For the sake of a concrete demonstration [368, Sect. 1.5], consider a general state in four-dimensional Hilbert space. It can be written as a vector in \mathbb{C}^4 , which can be parameterized by

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4)^T, \quad \text{with } \alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{C}, \quad (12.7)$$

and suppose (wrongly) (12.7) that all such states can be written in terms of a tensor product of two quasi-vectors in \mathbb{C}^2

$$(a_1, a_2)^T \otimes (b_1, b_2)^T \equiv (a_1 b_1, a_1 b_2, a_2 b_1, a_2 b_2)^T, \quad \text{with } a_1, a_2, b_1, b_2 \in \mathbb{C}. \quad (12.8)$$

A comparison of the coordinates in (12.7) and (12.8) yields

$$\alpha_1 = a_1 b_1, \quad \alpha_2 = a_1 b_2, \quad \alpha_3 = a_2 b_1, \quad \alpha_4 = a_2 b_2. \quad (12.9)$$

By taking the quotient of the two first and the two last equations, and by equating these quotients, one obtains

$$\frac{\alpha_1}{\alpha_2} = \frac{b_1}{b_2} = \frac{\alpha_3}{\alpha_4}, \text{ and thus } \alpha_1\alpha_4 = \alpha_2\alpha_3. \quad (12.10)$$

How can we imagine this? As in many cases, states in the Bell basis, and, in particular, the Bell state, serve as a sort of *Rosetta Stone* for an understanding of this quantum feature. The *Bell state* $|\Psi^-\rangle$ is a typical example of an entangled state; or, more generally, states in the *Bell basis* can be defined and, with $|0\rangle = (1, 0)^T$ and $|1\rangle = (0, 1)^T$, encoded by

$$|\Psi^\mp\rangle = \frac{1}{\sqrt{2}} (|01\rangle \mp |10\rangle) = \begin{pmatrix} 0 \\ 1 \\ \mp 1 \\ 0 \end{pmatrix}, |\Phi^\mp\rangle = \frac{1}{\sqrt{2}} (|00\rangle \mp |11\rangle) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \mp 1 \end{pmatrix}. \quad (12.11)$$

For instance, in the case of $|\Psi^-\rangle$ a comparison of coefficient yields

$$\begin{aligned} \alpha_1 = a_1b_1 = 0, \quad \alpha_2 = a_1b_2 = \frac{1}{\sqrt{2}}, \\ \alpha_3 = a_2b_1 = -\frac{1}{\sqrt{2}}, \quad \alpha_4 = a_2b_2 = 0; \end{aligned} \quad (12.12)$$

and thus entanglement, since

$$\alpha_1\alpha_4 = 0 \neq \alpha_2\alpha_3 = \frac{1}{2}. \quad (12.13)$$

This shows that $|\Psi^-\rangle$ cannot be considered as a two particle product state. Indeed, the state can only be characterized by considering the *relative properties* of the two particles – in the case of $|\Psi^-\rangle$ they are associated with the statements [588]: “the quantum numbers (in this case “0” and “1”) of the two particles are different in (at least) two orthogonal directions.”

The Bell basis symbolizing entanglement and non-individuality can, in an ad hoc manner, be generated from a non-entangled, individual state symbolized by elements of the Cartesian standard basis in 4-dimensional real space \mathbb{R}^4

$$|\mathbf{e}_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |\mathbf{e}_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\mathbf{e}_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |\mathbf{e}_4\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (12.14)$$

by arranging the coordinates (12.11) of the Bell basis as row or column vectors, thereby forming the respective unitary transformation

$$\begin{aligned}
\mathbf{U} &= |\Psi^-\rangle\langle\mathbf{e}_1| + |\Psi^+\rangle\langle\mathbf{e}_2| + |\Phi^-\rangle\langle\mathbf{e}_3| + |\Phi^+\rangle\langle\mathbf{e}_4| = \\
&= (|\Psi^-\rangle, |\Psi^+\rangle, |\Phi^-\rangle, |\Phi^+\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}. \quad (12.15)
\end{aligned}$$

Successive application of \mathbf{U} and its inverse \mathbf{U}^\top transforms an individual, non-entangled state from the Cartesian basis back and forth into an entangled, non-individual state from the Bell basis. For the sake of another demonstration, consider the following perfectly cyclic evolution which permutes all (non-)entangled states corresponding to the Cartesian and Bell bases:

$$|\mathbf{e}_1\rangle \xrightarrow{\mathbf{U}} |\Psi^-\rangle \xrightarrow{\mathbf{V}} |\mathbf{e}_2\rangle \xrightarrow{\mathbf{U}} |\Psi^+\rangle \xrightarrow{\mathbf{V}} |\mathbf{e}_3\rangle \xrightarrow{\mathbf{U}} |\Phi^-\rangle \xrightarrow{\mathbf{V}} |\mathbf{e}_4\rangle \xrightarrow{\mathbf{U}} |\Phi^+\rangle \xrightarrow{\mathbf{V}} |\mathbf{e}_1\rangle. \quad (12.16)$$

This evolution is facilitated by \mathbf{U} of Eq. (12.15), as well as by the following additional unitary transformation [460]:

$$\begin{aligned}
\mathbf{V} &= |\mathbf{e}_2\rangle\langle\Psi^-| + |\mathbf{e}_3\rangle\langle\Psi^+| + |\mathbf{e}_4\rangle\langle\Phi^-| + |\mathbf{e}_1\rangle\langle\Phi^+| = \\
&= \begin{pmatrix} \langle\Phi^+| \\ \langle\Psi^-| \\ \langle\Psi^+| \\ \langle\Phi^-| \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}. \quad (12.17)
\end{aligned}$$

One of the ways thinking of this kind of “breathing in and out of individuality and entanglement” is in terms of sampling and scrambling information, as quoted from Chiao [251, p. 27] (reprinted in [350]): “*Nothing has really been erased here, only scrambled!*” Indeed, mere re-coding or “scrambling,” and not erasure or creation of information, is tantamount to, and an expression and direct consequence of, the unitary evolution of the quantum state.

12.9 Quantum Probabilities

So far, quantum theory lacks probabilities. These will be introduced and compared to classical probabilities next. Indeed, for the sake of appreciating the novel features of quantum probabilities and correlations, as well as the (joint) expectations of quantum observables, a short excursion into classical probability theory is useful.

12.9.1 *Boole's Conditions of Possible Experience*

Already George Boole, although better known for his symbolic logic calculus of propositions aka *Laws of Thought* [66], pointed out that the probabilities of certain events, as well as their (joint) occurrence are subject to linear constraints [45–50, 66, 67, 163, 181–183, 221, 257, 258, 328, 421, 424, 524, 541–543]. A typical problem considered by Boole was the following [67, p. 229]: “Let p_1, p_2, \dots, p_n represent the probabilities given in the data. As these will in general not be the probabilities of unconnected events, they will be subject to other conditions than that of being positive proper fractions, . . . Those other conditions will, as will hereafter be shown, be capable of expression by equations or inequations reducible to the general form $a_1 p_1 + a_2 p_2 + \dots + a_n p_n + a \geq 0$, a_1, a_2, \dots, a_n, a being numerical constants which differ for the different conditions in question. These . . . may be termed the conditions of possible experience.”

Independently, Bell [40] derived some bounds on classical joint probabilities which relate to quantized systems insofar as they can be tested and falsified in the quantum regime by measuring subsets of compatible observables (possibly by Einstein–Podolsky–Rosen type [196] counterfactual inference) – one at a time – on different subensembles prepared in the same state. Thereby, in hindsight, it appears to be a bitter turn of history of thought that Bell, a staunch classical realist, who found wanting [41] previous attempts [552, 554], created one of the most powerful theorems used against (local) hidden variables. The present form of the “Bell inequalities” is due to Wigner [572] (cf. Sakurai [439, pp. 241–243] and Pitowsky [397, Footnote 13]. Fine [215] later pointed out that deterministic hidden variables just amount to suitable joint probability functions.

In referring to a later paper by Bell [42], Froissart [143, 227] proposed a general constructive method to produce all “maximal” (in the sense of tightest) constraints on classical probabilities and correlations for arbitrary physical configurations. This method uses all conceivable types of classical correlated outcomes, represented as matrices (or higher dimensional objects) which are the vertices [227, p. 243] “of a polyhedron which is their convex hull. Another way of describing this convex polyhedron is to view it as an intersection of half-spaces, each one corresponding to a face. The points of the polyhedron thus satisfy as many inequations as there are faces. Computation of the face equations is straightforward but tedious.” That is, certain “optimal” Bell-type inequalities can be interpreted as defining half-spaces (“below-above,” “inside-outside”) which represent the faces of a convex correlation polytope.

Later Pitowsky pointed out that any Bell-type inequality can be interpreted as Boole’s condition of possible experience [396–400, 407]. Pitowsky does not quote Froissart but mentions [396, p. 1556] that he had been motivated by a (series of) paper(s) by Garg and Mermin [235] (who incidentally did not mention Froissart either) on *Farkas’ Lemma*. Their concerns were linear constraints on pair distributions, derivable from the existence of higher-order distributions; constraints which turn out to be Bell-type inequalities; derivable as facets of convex correlation

polytopes. The Garg and Mermin paper is important because it concentrates on the “inverse” problem: rather than finding high-order distributions from low-order ones, they consider the question of whether or not those high-order distributions could return random variables with first order distributions as marginals. One of the examples mentioned [235, p. 2] are “three dichotomic variables each of which assumes either the value 1 or -1 with equal probability, and all the pair distributions vanish unless the members of the pair have different values, then any third-order distribution would have to vanish unless all three variables had different values. There can therefore be no third-order distribution.” (I mention this also because of the similarity with Specker’s parable of three boxes [479, 521].) A very similar question had also been pursued by Vorob’ev [556] and Kellerer [304, 305], who inspired Klyachko [312], as neither one of the previous authors are mentioned. [To be fair, in the reference section of an unpublished previous paper [311] Klyachko mentions Pitowsky two times; one reference not being cited in the main text.]

12.9.2 *Classical Strategies: Probabilities from Convex Sum of Truth Assignments and the Convex Polytope Method*

The gist of the classical strategy is to obtain all conceivable probabilities by a convex polytope method: any classical probability distribution can be written as a convex sum of all of the conceivable “extreme” cases. These “extreme” cases can be interpreted as classical truth assignments; or, equivalently, as two-valued states. A two-valued state is a function on the propositional structure of elementary observables, assigning any proposition the values “0” and “1” if they are (for a particular “extreme” case) “false” or “true,” respectively. “Extreme” cases are subject to criteria defined later in Sect. 12.9.4. The first explicit use [502, 506, 511, 521] (see Pykacz [423] for an early use of two-valued states) of the polytope method for deriving bounds using two-valued states on logics with intertwined contexts seems to have been for the pentagon logic, discussed in Sect. 12.9.8.3) and cat’s cradle logic (also called “Käfer,” the German word for “bug,” by Specker), discussed in Sect. 12.9.8.4.

More explicitly, suppose that there be as many, say, k , “weights” $\lambda_1, \dots, \lambda_k$ as there are two-valued states (or “extreme” cases, or truth assignments, if you prefer this denominations). Then convexity demands that all of these weights are positive and sum up to one; that is,

$$\begin{aligned} \lambda_1, \dots, \lambda_k &\geq 0, \text{ and} \\ \lambda_1 + \dots + \lambda_k &= 1. \end{aligned} \tag{12.18}$$

Suppose further that for any particular, say, the i th, two-valued state (or the i th “extreme” case, or the i th truth assignment, if you prefer this denomination), all the, say, m , “relevant” terms – relevance here merely means that we want them to contribute to the linear bounds denoted by Boole as conditions of possible experience,

as discussed in Sect. 12.9.6 – are “lumped” or combined together and identified as vector components of a vector $|\mathbf{x}_i\rangle$ in an m -dimensional vector space \mathbb{R}^m ; that is,

$$|\mathbf{x}_i\rangle = (x_{i_1}, x_{i_2}, \dots, x_{i_m})^\top. \quad (12.19)$$

Note that any particular convex [see Eq. (12.18)] combination

$$|\mathbf{w}(\lambda_1, \dots, \lambda_k)\rangle = \lambda_1 |\mathbf{x}_1\rangle + \dots + \lambda_k |\mathbf{x}_k\rangle \quad (12.20)$$

of the k weights $\lambda_1, \dots, \lambda_k$ yields a valid – that is consistent, subject to the criteria defined later in Sect. 12.9.4 – classical probability distribution, characterized by the vector $|\mathbf{w}(\lambda_1, \dots, \lambda_k)\rangle$. These k vectors $|\mathbf{x}_1\rangle, \dots, |\mathbf{x}_k\rangle$ can be identified with *vertices* or *extreme points* (which cannot be represented as convex combinations of other vertices or extreme points), associated with the k two-valued states (or “extreme” cases, or truth assignments). Let $V = \{|\mathbf{x}_1\rangle, \dots, |\mathbf{x}_k\rangle\}$ be the set of all such vertices.

For any such subset V (of vertices or extreme points) of \mathbb{R}^m , the *convex hull* is defined as the smallest convex set in \mathbb{R}^m containing V [230, Sect. 2.10, p. 6]. Based on its vertices a convex \mathcal{V} -polytope can be defined as the subset of \mathbb{R}^m which is the convex hull of a finite set of vertices or extreme points $V = \{|\mathbf{x}_1\rangle, \dots, |\mathbf{x}_k\rangle\}$ in \mathbb{R}^m :

$$\begin{aligned} P &= \text{Conv}(V) = \\ &= \left\{ \sum_{i=1}^k \lambda_i |\mathbf{x}_i\rangle \mid \lambda_1, \dots, \lambda_k \geq 0, \sum_{i=1}^k \lambda_i = 1, |\mathbf{x}_i\rangle \in V \right\}. \end{aligned} \quad (12.21)$$

A convex \mathcal{H} -polytope can also be defined as the intersection of a finite set of half-spaces, that is, the solution set of a finite system of n linear inequalities:

$$P = P(A, \mathbf{b}) = \left\{ |\mathbf{x}\rangle \in \mathbb{R}^m \mid \mathbf{A}_i |\mathbf{x}\rangle \leq |\mathbf{b}\rangle \text{ for } 1 \leq i \leq n \right\}, \quad (12.22)$$

with the condition that the set of solutions is bounded, such that there is a constant c such that $\| |\mathbf{x}\rangle \| \leq c$ holds for all $|\mathbf{x}\rangle \in P$. \mathbf{A}_i are matrices and $|\mathbf{b}\rangle$ are vectors with real components, respectively. Due to the Minkowski-Weyl “main” representation theorem [22, 230, 254, 274, 361, 449, 590] every \mathcal{V} -polytope has a description by a finite set of inequalities. Conversely, every \mathcal{H} -polytope is the convex hull of a finite set of points. Therefore the \mathcal{H} -polytope representation in terms of inequalities as well as the \mathcal{V} -polytope representation in terms of vertices, are equivalent, and the term convex polytope can be used for both and interchangeably. A k -dimensional convex polytope has a variety of *faces* which are again convex polytopes of various dimensions between 0 and $k - 1$. In particular, the 0-dimensional faces are called *vertices*, the 1-dimensional faces are called *edges*, and the $k - 1$ -dimensional faces are called *facets*.

The solution of the *hull problem*, or the *convex hull computation*, is the determination of the convex hull for a given finite set of k extreme points $V = \{|\mathbf{x}_1\rangle, \dots, |\mathbf{x}_k\rangle\}$

in \mathbb{R}^m (the general hull problem would also tolerate points inside the convex polytope); in particular, its representation as the intersection of half-spaces defining the facets of this polytope – serving as criteria of what lies “inside” and “outside” of the polytope – or, more precisely, as a set of solutions to a minimal system of linear inequalities. As long as the polytope has a non-empty interior and is full-dimensional (with respect to the vector space into which it is imbedded) there are only inequalities; otherwise, if the polytope lies on a hyperplane one obtains also equations.

For the sake of a familiar example, consider the regular 3-cube, which is the convex hull of the 8 vertices in \mathbb{R}^3 of $V = \{ (0, 0, 0)^\top, (0, 0, 1)^\top, (0, 1, 0)^\top, (1, 0, 0)^\top, (0, 1, 1)^\top, (1, 1, 0)^\top, (1, 0, 1)^\top, (1, 1, 1)^\top \}$. The cube has 8 vertices, 12 edges, and 6 facets. The half-spaces defining the regular 3-cube can be written in terms of the 6 facet inequalities $0 \leq x_1, x_2, x_3 \leq 1$.

Finally the *correlation polytope* can be defined as the convex hull of all the vertices or extreme points $|\mathbf{x}_1\rangle, \dots, |\mathbf{x}_k\rangle$ in V representing the (k per two-valued state) “relevant” terms evaluated for all the two-valued states (or “extreme” cases, or truth assignments); that is,

$$\begin{aligned} \text{Conv}(V) &= \left\{ |\mathbf{w}(\lambda_1, \dots, \lambda_k)\rangle \right\} \\ \left| \mathbf{w}(\lambda_1, \dots, \lambda_k)\rangle &= \lambda_1 |\mathbf{x}_1\rangle + \dots + \lambda_k |\mathbf{x}_k\rangle, \\ \lambda_1, \dots, \lambda_k &\geq 0, \quad \lambda_1 + \dots + \lambda_k = 1, \quad |\mathbf{x}_i\rangle \in V \end{aligned} \quad (12.23)$$

The convex \mathcal{H} -polytope – associated with the convex \mathcal{V} -polytope in (12.23) – which is the intersection of a finite number of half-spaces, can be identified with Boole’s conditions of possible experience.

A similar argument can be put forward for bounds on expectation values, as the expectations of dichotomic $E \in \{-1, +1\}$ -observables can be considered as affine transformations of two-valued states $v \in \{0, 1\}$; that is, $E = 2v - 1$. One might even imagine such bounds on arbitrary values of observables, as long as affine transformations are applied. Joint expectations from products of probabilities transform non-linearly, as, for instance $E_{12} = (2v_1 - 1)(2v_2 - 1) = 4v_1v_2 - 2(v_1 + v_2) - 1$. So, given some bounds on (joint) expectations; these can be translated into bounds on (joint) probabilities by substituting $2v_i - 1$ for expectations E_i . The converse is also true: bounds on (joint) probabilities can be translated into bounds on (joint) expectations by $v_i = (E_i + 1)/2$.

This method of finding classical bounds must fail if, such as for Kochen–Specker configurations, there are no or “too few” (such that there exist two or more atoms which cannot be distinguished by any two-valued state) two-valued states. In this case one may ease the assumptions; in particular, abandon admissibility, arriving at what has been called non-contextual inequalities [92].

12.9.3 Context and Greechie Orthogonality Diagrams

Henceforth a context will be any Boolean (sub-)algebra of experimentally observable propositions. The terms *block* or *classical mini-universe* will be used synonymously.

In classical physics there is only one context – and that is the entire set of observables. There exist models such as partition logics [184, 506, 511] – realizable by Wright’s generalized urn model [578] or automaton logic [444–446, 499], – which are still quasi-classical but have more than one, possibly intertwined, contexts. Two contexts are intertwined if they share one or more common elements. In what follows we shall only consider contexts which, if at all, intertwine at a single atomic proposition.

For such configurations Greechie has proposed a kind of *orthogonality diagram* [249, 300, 523] in which

1. entire contexts (Boolean subalgebras, blocks) are drawn as smooth lines, such as straight (unbroken) lines, circles or ellipses;
2. the atomic propositions of the context are drawn as circles; and
3. contexts intertwining at a single atomic proposition are represented as non-smoothly connected lines, broken at that proposition.

In Hilbert space realizations, the straight lines or smooth curves depicting contexts represent orthogonal bases (or, equivalently, maximal observables, Boolean subalgebras or blocks), and points on these straight lines or smooth curves represent elements of these bases; that is, two points on the same straight line or smooth curve represent two orthogonal basis elements. From dimension three onwards, bases may intertwine [240] by possessing common elements.

12.9.4 Two-Valued Measures, Frame Functions and Admissibility of Probabilities and Truth Assignments

In what follows we shall use notions of “truth assignments” on elements of logics which carry different names for related concepts:

1. The quantum logic community uses the term *two-valued state*; or, alternatively, *valuation* for a *total* function v on all elements of some logic L mapping $v : L \rightarrow [0, 1]$ such that [420, Definition 2.1.1, p. 20]
 - a. $v(\mathbb{I}) = 1$,
 - b. if $\{a_i, i \in \mathbb{N}\}$ is a sequence of mutually orthogonal elements in L – in particular, this applies to atoms within the same context (block, Boolean subalgebra) – then the two-valued state is additive on those elements a_i ; that is, additivity holds:

$$v\left(\bigvee_{i \in \mathbb{N}}\right) = \sum_{i \in \mathbb{N}} v(a_i). \quad (12.24)$$

2. Gleason has used the term *frame function* [240, p. 886] of weight 1 for a separable Hilbert space \mathfrak{H} as a total, real-valued (not necessarily two-valued) function f defined on the (surface of the) unit sphere of \mathfrak{H} such that if $\{a_i, i \in \mathbb{N}\}$ represents an orthonormal basis of \mathfrak{H} , then additivity

$$\sum_{i \in \mathbb{N}} f(a_i) = 1. \quad (12.25)$$

holds for all orthonormal bases (contexts, blocks) of the logic based on \mathfrak{H} .

3. A dichotomic total function $v : L \rightarrow [0, 1]$ will be called *strongly admissible* if
- a. within every context $C = \{a_i, i \in \mathbb{N}\}$, a single atom a_j is assigned the value one: $v(a_j) = 1$; and
 - b. all other atoms in that context are assigned the value zero: $v(a_i \neq a_j) = 0$. Physically this amounts to only one elementary proposition being true; the rest of them are false. (One may think of an array of mutually exclusively firing detectors.)
 - c. Non-contextuality, stated explicitly]: The value of any observable, and, in particular, of an atom in which two contexts intertwine, does not depend on the context. It is context-independent.
4. In order to cope with value indefiniteness (cf. Sect. 12.9.8.7), a weaker form of admissibility has been proposed [3–6] which is no total function but rather is a *partial* function which may remain undefined (indefinite) on some elements of L : A dichotomic partial function $v : L \rightarrow [0, 1]$ will be called *admissible* if the following two conditions hold for every context C of L :
- a. if there exists a $a \in C$ with $v(a) = 1$, then $v(b) = 0$ for all $b \in C \setminus \{a\}$;
 - b. if there exists a $a \in C$ with $v(b) = 0$ for all $b \in C \setminus \{a\}$, then $v(a) = 1$;
 - c. the value assignments of all other elements of the logic not covered by, if necessary, successive application of the admissibility rules, are undefined and thus the atom remains value indefinite.

Unless otherwise mentioned (such as for contextual value assignments or admissibility discussed in Sect. 12.9.8.7) the quantum logical (I), Gleason type (II), strong admissibility (III) notions of two-valued states will be used. Such two valued states (probability measures) are interpretable as (pre-existing) truth assignments; they are sometimes also referred to as a *Kochen–Specker value assignment* [583].

12.9.5 Why Classical Correlation Polytopes?

A *caveat* seems to be in order from the very beginning: in what follows correlation polytopes arise from classical (and quasi-classical) situations. The considerations are relevant for quantum mechanics only insofar as the quantum probabilities could violate classical bounds; that is, if the quantum tests violate those bounds by “lying outside” of the classical correlation polytope.

There exist at least two good reasons to consider (correlation) polytopes for bounds on classical probabilities, correlations and expectation values:

1. they represent a systematic way of enumerating the probability distributions and deriving constraints – Boole’s conditions of possible experience – on them;
2. one can be sure that these constraints and bounds are *optimal* in the sense that they are guaranteed to yield inequalities which are best criteria for classicality.

It is not evident to see why, with the methods by which they have been obtained, Bell’s original inequality [41, 42] or the Clauser–Horne–Shimony–Holt inequality [145] should be “optimal” at the time they were presented. Their derivation involve estimates which appear ad hoc; and it is not immediately obvious that bounds based on these estimates could not be improved. The correlation polytope method, on the other hand, offers a conceptually clear framework for a derivation of all classical bounds on higher-order distributions.

12.9.6 What Terms May Enter Classical Correlation Polytopes?

What can enter as terms in such correlation polytopes? To quote Pitowsky [397, p. 38], “Consider n events A_1, A_2, \dots, A_n , in a classical event space ... Denote $p_i = \text{probability}(A_i)$, $p_{ij} = \text{probability}(A_i \cap A_j)$, and more generally $p_{i_1 i_2 \dots i_k} = \text{probability}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})$, whenever $1 \leq i_1 < i_2 < \dots < i_k \leq n$. We assume no particular relations among the events. Thus A_1, \dots, A_n are not necessarily distinct, they can be dependent or independent, disjoint or non-disjoint etc.”

However, although the events A_1, \dots, A_n may be in any relation to one another, one has to make sure that the respective probabilities, and, in particular, the extreme cases – the two-valued states interpretable as truth assignments – properly encode the logical or empirical relations among events. In particular, when it comes to an enumeration of cases, consistency must be retained. For example, suppose one considers the following three propositions: A_1 : “it rains in Vienna,” A_3 : “it rains in Vienna or it rains in Auckland.” It cannot be that A_2 is less likely than A_1 ; therefore, the two-valued states interpretable as truth assignments must obey $p(A_2) \geq p(A_1)$, and in particular, if A_1 is true, A_2 must be true as well. (It may happen though that A_1 is false while A_2 is true.) Also, mutually exclusive events cannot be true simultaneously.

These admissibility and consistency requirements are considerably softened in the case of non-contextual inequalities [92], where subclassicality – the requirement that among a complete (maximal) set of mutually exclusive observables only one is true and all others are false (equivalent to one important criterion for Gleason’s frame function [240]) – is abandoned. To put it pointedly, in such scenarios, the simultaneous existence of inconsistent events such as A_1 : “it rains in Vienna,” A_2 : “it does not rain in Vienna” are allowed; that is, $p(\text{“it rains in Vienna”}) = p(\text{“it does not rain in Vienna”}) = 1$. The reason for this rather desperate step is that, for Kochen–Specker type configurations, there are no classical truth assignments satisfying the classical admissibility rules; therefore the latter are abandoned. (With the admissibility rules goes the classical Kolmogorovian probability axioms even within classical Boolean subalgebras.)

It is no coincidence that most calculations are limited – or rather limit themselves because there is no formal reasons to go to higher orders – to the joint probabilities or expectations of just two observables: there is no easy “workaround” of quantum complementarity. The Einstein–Podolsky–Rosen setup [196] offers one for just two complementary contexts at the price of counterfactuals, but there seems to be no generalization to three or more complementary contexts in sight [448].

12.9.7 General Framework for Computing Boole’s Conditions of Possible Experience

As pointed out earlier, Froissart and Pitowsky, among others such as Tsirelson, have sketched a very precise algorithmic framework for constructively finding all conditions of possible experience. In particular, Pitowsky’s later method [397–400, 407], with slight modifications for very general non-distributive propositional structures such as the pentagon logic [506, 511, 521], goes like this:

1. define the terms which should enter the bounds;
2.
 - a. if the bounds should be on the probabilities: evaluate all two-valued measures interpretable as truth assignments;
 - b. if the bounds should be on the expectations: evaluate all value assignments of the observables;
 - c. if (as for non-contextual inequalities) the bounds should be on some pre-defined quantities: evaluate all value definite pre-assigned quantities;
3. arrange these terms into vectors whose components are all evaluated for a fixed two-valued state, one state at a time; one vector per two-valued state (truth assignment), or (for expectations) per value assignments of the observables, or (for non-contextual inequalities) per value-assignment;
4. consider the set of all obtained vectors as vertices of a convex polytope;
5. solve the convex hull problem by computing the convex hull, thereby finding the smallest convex polytope containing all these vertices. The solution can be represented as the half-spaces (characterizing the facets of the polytope) formalized

by (in)equalities – (in)equalities which can be identified with Boole’s conditions of possible experience.

Froissart [227] and Tsirelson [143] are not much different; they arrange joint probabilities for two random variables into matrices instead of “delineating” them as vectors; but this difference is notational only. We shall explicitly apply the method to various configurations next.

12.9.8 Some Examples

In what follows we shall enumerate several (non-)trivial – that is, non-Boolean in the sense of pastings [249, 300, 376, 420] of Boolean subalgebras. Suppose some points or vertices in \mathbb{R}^n are given. The convex hull problem of finding the smallest convex polytope containing all these points or vertices, given the latter, will be solved evaluated with Fukuda’s *cddlib package cddlib-094h* [229] (using GMP [223]) implementing the *double description method* [22, 23, 231].

12.9.8.1 Trivial Cases

Bounds on the Probability of One Observable

The case of a single variable has two extreme cases: false $\equiv 0$ and true $\equiv 1$, resulting in the two vertices (0) as well as (1) , respectively. The corresponding hull problem yields a probability “below 0” as well as “above 1,” respectively; thus solution this rather trivial hull problem yields $0 \leq p_1 \leq 1$. For dichotomic expectation values ± 1 a similar argument yields $-1 \leq E_1 \leq 1$.

Bounds on the (Joint) Probabilities and Expectations of Two Observables

The next trivial case is just two dichotomic (two values) observables and their joint probability. The respective logic is generated by the pairs (overline indicates negation) $a_1a_2, a_1\bar{a}_2, \bar{a}_1a_2, \bar{a}_1\bar{a}_2$, representable by a single Boolean algebra 2^4 , whose atoms are these pairs: $a_1a_2, a_1\bar{a}_2, \bar{a}_1a_2, \bar{a}_1\bar{a}_2$. For single Boolean algebras with k atoms, there are k two-valued measures; in this case $k = 4$.

For didactic purposes this case has been covered *ad nauseam* in Pitowsky’s introductions [396–400, 407]; so it is just mentioned without further discussion: take the probabilities two observables p_1 and p_2 , and a their joint variable p_{12} and “bundle” them together into a vector $(p_1, p_2, p_1 \wedge p_2 \equiv p_{12} = p_1 p_2)^T$ of three-dimensional vector space. Then enumerate all four extreme cases – the two-valued states interpretable as truth assignments – involving two observables p_1 and p_2 , and a their joint variable p_{12} very explicitly false-false-false, false-true-false, true-false-false, and true-true-true, or by numerical encoding, 0-0-0, 0, 1, 0, 1, 0, 0, and 1-1-1, yielding the four vectors

$$\begin{aligned} |v_1\rangle &= (0, 0, 0)^\top, |v_2\rangle = (0, 1, 0)^\top, \\ |v_3\rangle &= (1, 0, 0)^\top, |v_4\rangle = (1, 1, 1)^\top. \end{aligned} \quad (12.26)$$

Solution of the hull problem for the polytope

$$\left\{ \begin{aligned} & \lambda_1 |v_1\rangle + \lambda_2 |v_2\rangle + \lambda_3 |v_3\rangle + \lambda_4 |v_4\rangle \\ & \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1, \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0 \end{aligned} \right\} \quad (12.27)$$

yields the “inside-outside” inequalities of the half-spaces corresponding to the four facets of this polytope:

$$\begin{aligned} p_1 + p_2 - p_{12} &\leq 1, \\ 0 &\leq p_{12} \leq p_1, p_2. \end{aligned} \quad (12.28)$$

For the expectation values of two dichotomic observables ± 1 a similar argument yields

$$\begin{aligned} E_1 + E_2 - E_{12} &\leq 1, \\ -E_1 + E_2 + E_{12} &\leq 1, \\ E_1 - E_2 + E_{12} &\leq 1, \\ -E_1 - E_2 - E_{12} &\leq 1. \end{aligned} \quad (12.29)$$

Bounds on the (Joint) Probabilities and Expectations of Three Observables

Very similar calculations, taking into account three observables and their joint probabilities and expectations, yield

$$\begin{aligned} p_1 + p_2 + p_3 - p_{12} - p_{13} - p_{23} + p_{123} &\leq 1, \\ -p_1 + p_{12} + p_{13} - p_{123} &\leq 0, \\ -p_2 + p_{12} + p_{23} - p_{123} &\leq 0, \\ -p_3 + p_{13} + p_{23} - p_{123} &\leq 0, \\ p_{12}, p_{13}, p_{23} &\geq p_{123} \geq 0. \end{aligned} \quad (12.30)$$

and

$$\begin{aligned} -E_{12} - E_{13} - E_{23} &\leq 1 \\ -E_{12} + E_{13} + E_{23} &\leq 1, \\ E_{12} - E_{13} + E_{23} &\leq 1, \\ E_{12} + E_{13} - E_{23} &\leq 1, \\ -1 &\leq E_{123} \leq 1. \end{aligned} \quad (12.31)$$

12.9.8.2 Einstein–Podolsky–Rosen Type “Explosion” Setups of Joint Distributions Without Intertwined Contexts

The first non-trivial (in the sense that quantum probabilities and expectations violate the classical bounds) instance occurs for four observables in an Einstein–Podolski–Rosen type “explosion” setup [196], where n observables are measured on both sides, respectively.

Clauser–Horne–Shimony–Holt Case: 2 Observers, 2 Measurement Configurations per Observer

If just two observables are measured on the two sides, the facets of the polytope are the Bell–Wigner–Fine (in the probabilistic version) as well as the Clauser–Horne–Shimony–Holt (for joint expectations) inequalities; that is, for instance,

$$\begin{aligned} 0 \leq p_1 + p_4 - p_{13} - p_{14} + p_{23} - p_{24} \leq 1, \\ -2 \leq E_{13} + E_{14} + E_{23} - E_{24} \leq 2. \end{aligned} \quad (12.32)$$

To obtain a feeling, Fig. 12.1a depicts the Greechie orthogonality diagram of the 2 particle 2 observables per particle situation. Figure 12.1b enumerates all two-valued states thereon.

At this point it might be interesting to see how exactly the approach of Froissart and Tsirelson blends in [143, 227]. The only difference to the Pitowsky method – which enumerates the (two particle) correlations and expectations as *vector components* – is that Froissart and later and Tsirelson arrange the two-particle correlations and expectations as *matrix components*; so both differ only by notation. For instance, Froissart explicitly mentions [227, pp. 242–243] 10 extremal configurations of the two-particle correlations, associated with 10 matrices

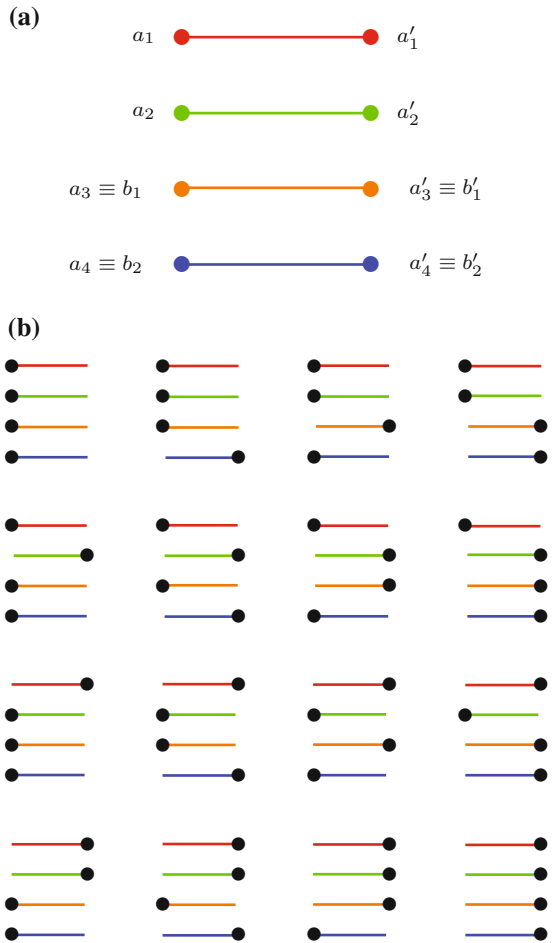
$$\begin{pmatrix} p_{13} = p_1 p_3 & p_{14} = p_1 p_4 \\ p_{23} = p_2 p_3 & p_{24} = p_2 p_4 \end{pmatrix} \quad (12.33)$$

containing 0s and 1s (the indices “1, 2” and “3, 4” are associated with the two sides of the Einstein–Podolsky–Rosen “explosion”-type setup, respectively), arranged in Pitowsky’s case as vector

$$(p_{13} = p_1 p_3, p_{14} = p_1 p_4, p_{23} = p_2 p_3, p_{24} = p_2 p_4). \quad (12.34)$$

For probability correlations the number of different matrices or vectors is 10 (and not 16 as could be expected from the 16 two-valued measures), since, as enumerated in Table 12.1 some such measures yield identical results on the two-particle correlations; in particular, $v_1, v_2, v_3, v_4, v_5, v_9, v_{13}$ yield identical matrices (in the Froissart case) or vectors (in the Pitowsky case).

Fig. 12.1 **a** Four contexts $\{a_1, a'_1\}$, $\{a_2, a'_2\}$ on one side, and $\{a_3 \equiv b_1, a'_3 \equiv b'_1\}$, $\{a_4 \equiv b_2, a'_4 \equiv b'_2\}$ on the other side of the Einstein–Podolsky–Rosen “explosion”–type setup are relevant for a computation of the Bell–Wigner–Fine (in the probabilistic version) as well as the Clauser–Horne–Shimony–Holt (for joint expectations) inequalities; **b** the 2^4 two-valued measures thereon, tabulated in Table 12.1, which are used to compute the vertices of the correlation polytopes. Full circles indicate the value “1 \equiv true”



Beyond the Clauser–Horne–Shimony–Holt Case: 2 Observers, More Measurement Configurations per Observer

The calculation for the facet inequalities for two observers and three measurement configurations per observer is straightforward and yields 684 inequalities [148, 407, 469]. If one considers (joint) expectations one arrives at novel ones which are not of the Clauser–Horne–Shimony–Holt type; for instance [469, p. 166, Eq. (4)],

Table 12.1 The 16 two-valued states on the 2 particle two observables per particle configuration, as drawn in Fig. 12.1b. Two-particle correlations appear green. There are 10 different such configurations, painted in red

#	a ₁	a ₂	a ₃	a ₄	a ₁₃	a ₁₄	a ₂₃	a ₂₄
v ₁	0	0	0	0	0	0	0	0
v ₂	0	0	0	1	0	0	0	0
v ₃	0	0	1	0	0	0	0	0
v ₄	0	0	1	1	0	0	0	0
v ₅	0	1	0	0	0	0	0	0
v ₆	0	1	0	1	0	0	0	1
v ₇	0	1	1	0	0	0	1	0
v ₈	0	1	1	1	0	0	1	1
v ₉	1	0	0	0	0	0	0	0
v ₁₀	1	0	0	1	0	1	0	0
v ₁₁	1	0	1	0	1	0	0	0
v ₁₂	1	0	1	1	1	1	0	0
v ₁₃	1	1	0	0	0	0	0	0
v ₁₄	1	1	0	1	0	1	0	1
v ₁₅	1	1	1	0	1	0	1	0
v ₁₆	1	1	1	1	1	1	1	1

$$\begin{aligned}
 -4 \leq & -E_2 + E_3 - E_4 - E_5 + E_{14} - E_{15} + \\
 & + E_{24} + E_{25} + E_{26} - E_{34} - E_{35} + E_{36}, \\
 -4 \leq & E_1 + E_2 + E_4 + E_5 + E_{14} + E_{15} + \\
 & + E_{16} + E_{24} + E_{25} - E_{26} + E_{34} - E_{35}.
 \end{aligned}
 \tag{12.35}$$

As already mentioned earlier, these bounds on classical expectations [469] translate into bounds on classical probabilities [148, 407] (and *vice versa*) if the affine transformations $E_i = 2v_i - 1$ [and conversely $v_i = (E_i + 1)/2$] are applied.

Here a word of warning is in order: if one only evaluates the vertices from the joint expectations (and not also the single particle expectations), one never arrives at the novel inequalities of the type listed in Eq. (12.35), but obtains 90 facet inequalities; among them 72 instances of the Clauser–Horne–Shimony–Holt inequality form, such as

$$\begin{aligned}
 E_{25} + E_{26} + E_{35} - E_{36} & \leq 2, \\
 E_{14} + E_{15} + E_{24} - E_{25} & \leq 2, \\
 -E_{25} - E_{26} - E_{35} + E_{36} & \leq 2, \\
 -E_{14} - E_{15} - E_{24} + E_{25} & \leq 2.
 \end{aligned}
 \tag{12.36}$$

They can be combined to yield (see also Ref. [469, p. 166, Eq. (4)])

$$-4 \leq E_{14} + E_{15} + E_{24} + E_{26} + E_{35} - E_{36} \leq 4.
 \tag{12.37}$$

For the general case of n qubits, algebraic methods different than the hull problem for polytopes have been suggested in Refs. [404, 443, 567, 594].

12.9.8.3 Intertwined Contexts

In the following we shall present a series of logics whose contexts (representable by maximal observables, Boolean subalgebras, blocks, or orthogonal bases) are intertwined; but “not much:” by assumption and for convenience, contexts intertwine in only one element; it does not happen that two contexts are pasted [249, 300, 376, 420] along two or more atoms. (They nevertheless might be totally identical.) Such intertwines – connecting contexts by pasting them together – can only occur from Hilbert space dimension three onwards, as contexts in lower-dimensional spaces cannot have the same element unless they are identical.

In Sect. 12.9.8.3 we shall first study the “firefly case” with just two contexts intertwined in one atom; then, in Sect. 12.9.8.3, proceed to the pentagon configuration with five contexts intertwined cyclically, then, in Sect. 12.9.8.4, paste two such pentagon logics to form a cat’s cradle (or, by another term, Specker’s bug) logic; and finally, in Sect. 12.9.8.6, connect two Specker bugs to arrive at a logic which has a so “meagre” set of states that it can no longer separate two atoms. As pointed out already by Kochen and Specker [314, p. 70,] this is no longer imbeddable into some Boolean algebra. It thus cannot be represented by a partition logic; and thus has neither any generalized urn and finite automata models nor classical probabilities separating different events. The case of logics allowing no two valued states will be covered consecutively.

Firefly Logic

Cohen presented [147, pp. 21–22] a classical realization of the first logic with just two contexts and one intertwining atom: a firefly in a box, observed from two sides of this box which are divided into two windows; assuming the possibility that sometimes the firefly does not shine at all. This firefly logic, which is sometimes also denoted by L_{12} because it has 12 elements (in a Hasse diagram) and 5 atoms, with the contexts defined by $\{a_1, a_2, a_5\}$ and $\{a_3, a_4, a_5\}$ is depicted in Fig. 12.2.

The five two-valued states on the firefly logic are enumerated in Table 12.2 and depicted in Fig. 12.3.

These two-valued states induce [506] a partition logic realization [184, 511] $\{\{\{1\}, \{2, 3\}, \{4, 5\}\}, \{\{1\}, \{2, 5\}, \{3, 4\}\}\}$ which in turn induce all classical probability

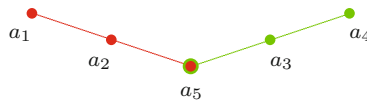


Fig. 12.2 Firefly logic with two contexts $\{a_1, a_2, a_5\}$ and $\{a_3, a_4, a_5\}$ intertwined in a_5

Table 12.2 Two-valued states on the firefly logic

#	a_1	a_2	a_3	a_4	a_5
v_1	0	0	0	0	1
v_2	0	1	0	1	0
v_3	0	1	1	0	0
v_4	1	0	1	0	0
v_5	1	0	0	1	0

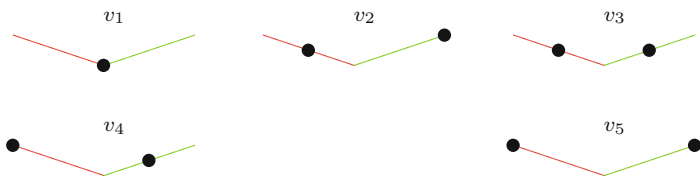


Fig. 12.3 Two-valued measures on the firefly logic. Filled circles indicate the value “1” interpretable as “true”

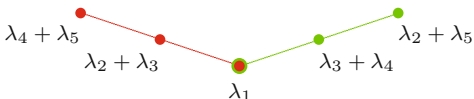


Fig. 12.4 Classical probabilities on the firefly logic with two contexts, as induced by the two-valued states, and subject to $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 = 1, 0 \leq \lambda_1, \dots, \lambda_5 \leq 1$

distributions, as depicted in Fig. 12.4. No representation in \mathbb{R}^3 is given here; but this is straightforward (just two orthogonal tripods with one identical leg), or can be read off from logics containing more such intertwined fireflies; such as in Fig. 12.6.

Pentagon Logic

Admissibility of two-valued states imposes conditions and restrictions on the two-valued states already for a single context (Boolean subalgebra): if one atom is assigned the value 1, all other atoms have to have value assignment(s) 0. This is even more so for intertwining contexts. For the sake of an example, consider two firefly logics pasted along an entire block, as depicted in Fig. 12.5. For such a logic we can state a “true-and-true implies true” rule: if the two-valued measure at the “outer extremities” is 1, then it must be 1 at its center atom.

We shall pursue this path of ever increasing restrictions through construction of pasted; that is, intertwined, contexts. This ultimately yields to non-classical logics which have no separating sets of two-valued states; and even, as in Kochen–Specker type configurations, to logics which do not allow for any two valued state interpretable as preassigned truth assignments.

Let us proceed by pasting more firefly logics together in “closed circles.” The next possibilities – two firefly logics forming either a triangle or a square Greechie orthogonal diagram – have no realization in three dimensional Hilbert space. The

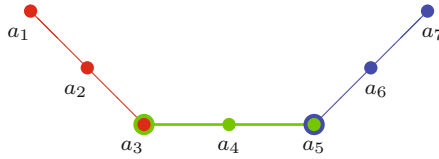


Fig. 12.5 Two firefly logics pasted along an entire context $\{a_3, a_4, a_5\}$ with the following property: if a two valued state v is $v(a_1) = v(a_6) = 1$, or $v(a_1) = v(a_7) = 1$, or $v(a_2) = v(a_6) = 1$, or $v(a_2) = v(a_7) = 1$, or then the “central atom” a_4 must be $v(a_4) = 1$. No representation in \mathbb{R}^3 is given here; but this is straightforward; or can be read off from logics containing more such intertwined fireflies; such as in Fig. 12.6

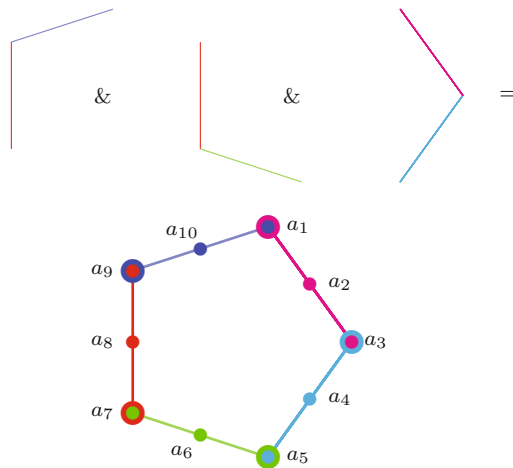


Fig. 12.6 Orthogonality diagram of the pentagon logic, which is a pasting of 3 firefly logics (two of which share an entire context), resulting in a pasting of five intertwined contexts $a = \{a_1, a_2, a_3\}$, $b = \{a_3, a_4, a_5\}$, $c = \{a_5, a_6, a_7\}$, $d = \{a_7, a_8, a_9\}$, $e = \{a_9, a_{10}, a_1\}$. They have a (quantum) realization in \mathbb{R}^3 consisting of the 10 projections associated with the one dimensional subspaces spanned by the vectors from the origin $(0, 0, 0)^T$ to $a_1 = (\sqrt[4]{5}, -\sqrt{\sqrt{5}-2}, \sqrt{2})^T$, $a_2 = (-\sqrt[4]{5}, -\sqrt{2+\sqrt{5}}, \sqrt{3-\sqrt{5}})^T$, $a_3 = (-\sqrt[4]{5}, \sqrt{2+\sqrt{5}}, \sqrt{3+\sqrt{5}})^T$, $a_4 = (\sqrt{5+\sqrt{5}}, \sqrt{3-\sqrt{5}}, 2\sqrt{-2+\sqrt{5}})^T$, $a_5 = (0, -\sqrt{\sqrt{5}-1}, 1)^T$, $a_6 = (-\sqrt{5+\sqrt{5}}, \sqrt{3-\sqrt{5}}, 2\sqrt{\sqrt{5}-2})^T$, $a_7 = (\sqrt[4]{5}, \sqrt{2+\sqrt{5}}, \sqrt{3+\sqrt{5}})^T$, $a_8 = (\sqrt[4]{5}, -\sqrt{2+\sqrt{5}}, \sqrt{3-\sqrt{5}})^T$, $a_9 = (-\sqrt[4]{5}, -\sqrt{\sqrt{5}-2}, \sqrt{2})^T$, $a_{10} = (0, \sqrt{2}, \sqrt{\sqrt{5}-2})^T$, respectively [523, Fig. 8, p. 5393]. Another such realization is $a_1 = (1, 0, 0)^T$, $a_2 = (0, 1, 0)^T$, $a_3 = (0, 0, 1)^T$, $a_4 = (1, -1, 0)^T$, $a_5 = (1, 1, 0)^T$, $a_6 = (1, -1, 2)^T$, $a_7 = (-1, 1, 1)^T$, $a_8 = (2, 1, 1)^T$, $a_9 = (0, 1, -1)^T$, $a_{10} = (0, 1, 1)^T$, respectively [532]

Table 12.3 Two-valued states on the pentagon

#	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
v_1	1	0	0	1	0	1	0	1	0	0
v_2	1	0	0	0	1	0	0	1	0	0
v_3	1	0	0	1	0	0	1	0	0	0
v_4	0	0	1	0	0	1	0	1	0	1
v_5	0	0	1	0	0	0	1	0	0	1
v_6	0	0	1	0	0	1	0	0	1	0
v_7	0	1	0	0	1	0	0	1	0	1
v_8	0	1	0	0	1	0	0	0	1	0
v_9	0	1	0	1	0	0	1	0	0	1
v_{10}	0	1	0	1	0	1	0	0	1	0
v_{11}	0	1	0	1	0	1	0	1	0	1
v_e	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0

next diagram realizably is obtained by a pasting of three firefly logics. It is the pentagon logic (also denoted as *orthomodular house* [300, p. 46, Fig. 4.4] and discussed in Ref. [50]; see also Birkhoff’s distributivity criterion [57, p. 90, Theorem 33], stating that, in particular, if some lattice contains a pentagon as sublattice, then it is not distributive [60]) which is subject to an old debate on “exotic” probability measures [577]. In terms of Greechie orthogonality diagrams there are two equivalent representations of the pentagon logic: one as a pentagon, as depicted [521] in Fig. 12.6 and one as a pentagram; thereby the indices of the intertwining edges (the non-intertwining ones follow suit) are permuted as follows: $1 \mapsto 1, 9 \mapsto 5, 7 \mapsto 9, 5 \mapsto 3, 3 \mapsto 7$. From a Greechie orthogonality point of view the pentagon representation is preferable over the pentagram, because the latter, although appearing more “magic,” might suggest the illusion that there are more intertwining contexts and observables as there actually are.

As pointed out by Wright [577, p 268] the pentagon has 11 “ordinary” two-valued states v_1, \dots, v_{11} , and one “exotic” dispersionless state v_e , which was shown by Wright to have neither a classical nor a quantum interpretation; all defined on the 10 atoms a_1, \dots, a_{10} . They are enumerated in Table 12.3. and depicted in Fig. 12.7.

These two-valued states directly translate into the classical probabilities depicted in Fig. 12.8.

The pentagon logic has quasi-classical realizations in terms of partition logics [184, 506, 511], such as generalized urn models [577, 578] or automaton logics [444–446, 499]. An early realization in terms of three-dimensional (quantum) Hilbert space can, for instance, be found in Ref. [523, pp. 5392–5393]; other such parametrizations are discussed in Refs. [24, 85, 86, 312].

The full hull problem, including all joint expectations of dichotomic ± 1 observables yields 64 inequalities enumerated in the supplementary material; among them

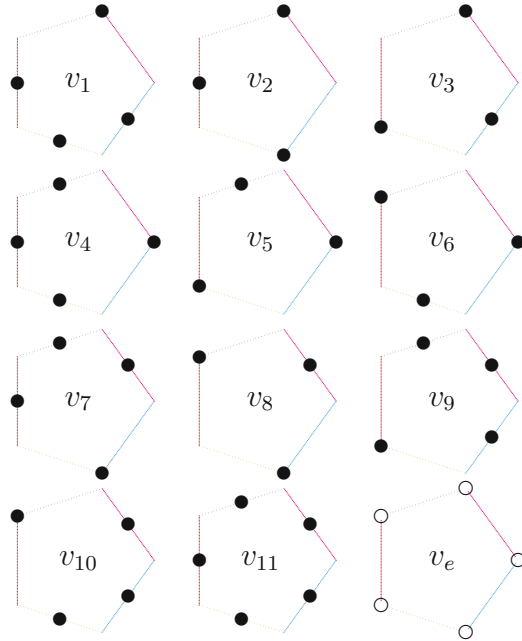


Fig. 12.7 Two-valued measures on the pentagon logic. Filled circles indicate the value “1” interpretable as “true.” In the last diagram non-filled circles indicate the value “ $\frac{1}{2}$ ”

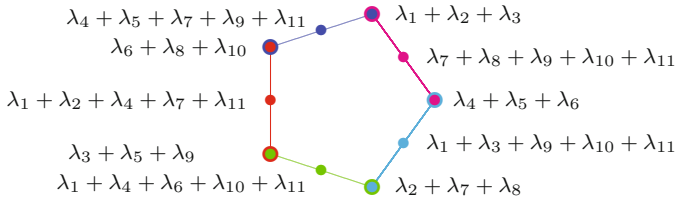


Fig. 12.8 Classical probabilities on the pentagon logic, $\lambda_1 + \dots + \lambda_{11} = 1$, $\lambda_1, \dots, \lambda_{11} \geq 0$, taken from Ref. [521]

$$\begin{aligned}
 E_{12} &\leq E_{45}, \quad E_{18} \leq E_{7,10}, \\
 E_{16} + E_{26} + E_{36} + E_{48} &\leq E_{18} + E_{28} + E_{34} + E_{59}, \\
 E_{14} + E_{18} + E_{28} &\leq 1 + E_{12} + E_{16} + E_{26} + E_{36} + E_{48} + E_{5,10}.
 \end{aligned}
 \tag{12.38}$$

The full hull computations for the probabilities p_1, \dots, p_{10} on all atoms a_1, \dots, a_{10} reduces to 16 inequalities, among them

$$\begin{aligned} +p_4 + p_8 + p_9 &\geq +p_1 + p_2 + p_6, \\ 2p_1 + p_2 + p_6 + p_{10} &\geq 1 + p_4 + p_8. \end{aligned} \quad (12.39)$$

If one considers only the five probabilities on the intertwining atoms, then the Bub-Stairs inequality $p_1 + p_3 + p_5 + p_7 + p_9 \leq 2$ results [24, 85, 86]. Concentration on the four non-intertwining atoms yields $p_2 + p_4 + p_6 + p_8 + p_{10} \geq 1$. Limiting the hull computation to adjacent pair expectations of dichotomic ± 1 observables yields the Klyachko–Can–Binicioglu–Shumovsky inequality [312]

$$E_{13} + E_{35} + E_{57} + E_{79} + E_{91} \geq 3. \quad (12.40)$$

12.9.8.4 Combo of Two Intertwined Pentagon Logics Forming a Specker Bug (or Pitowsky Cat’s Cradle) “True Implies False” Logic

The pasting of two pentagon logics results in ever tighter conditions for two-valued measures and thus truth value assignments: consider the Greechie orthogonality diagram of a logic drawn in Fig. 12.9. Specker [481] called this the “Käfer” (bug) Logic because of the similar shape with a bug. It has been introduced in 1963(5) by Kochen and Specker [313, Fig. 1, p. 182]; and subsequently used as a subset of the diagrams Γ_1 , Γ_2 and Γ_3 demonstrating the existence of quantum propositional structures with the “true implies true” property (cf. Sect. 12.9.8.5), the non-existence of any two-valued state (cf. Sect. 12.9.8.7), and the existence of a non-separating set of two-valued states (cf. Sect. 12.9.8.6), respectively [314].

Pitowsky called it (part of [429]) “cat’s cradle” [403, 405] (see also Refs. [39, Fig. B.1. p. 64], [483, pp. 588–589], [1, Sect. IV, Fig. 2] and [420, p. 39, Fig. 2.4.6] for early discussions). A partition logic, as well as a Hilbert space realization can be found in Refs. [511, 523]. There are 14 two-valued states which are listed in Table 12.4.

As already Pták and Pulmannová [420, p. 39, Fig. 2.4.6] as well as Pitowsky [403, 405] have pointed out, the reduction of some probabilities of atoms at intertwined contexts yields [521, p. 285, Eq. (11.2)]

$$p_1 + p_7 = \frac{3}{2} - \frac{1}{2}(p_{12} + p_{13} + p_2 + p_6 + p_8) \leq \frac{3}{2}. \quad (12.41)$$

A better approximation comes from the explicit parameterization of the classical probabilities on the atoms a_1 and a_7 , derivable from all the mutually disjoint two-valued states which do not vanish on those atoms, as depicted in Fig. 12.10: $p_1 = \lambda_1 + \lambda_2 + \lambda_3$, and $p_7 = \lambda_7 + \lambda_{10} + \lambda_{13}$. Because of additivity the 14 positive weights $\lambda_1, \dots, \lambda_{14} \geq 0$ must add up to 1; that is, $\sum_{i=1}^{14} \lambda_i = 1$. Therefore,

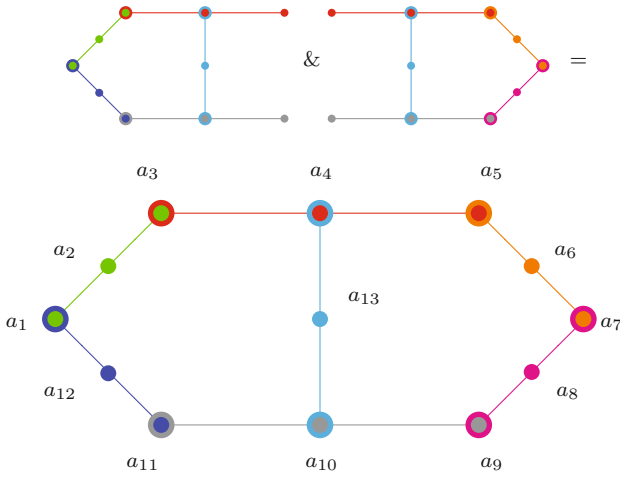


Fig. 12.9 Greechie diagram of the Specker bug (cat’s cradle) logic which results from a pasting of two pentagon logics sharing three common contexts. It is a pasting of seven intertwined contexts $a = \{a_1, a_2, a_3\}$, $b = \{a_3, a_4, a_5\}$, $c = \{a_5, a_6, a_7\}$, $d = \{a_7, a_8, a_9\}$, $e = \{a_9, a_{10}, a_{11}\}$, $f = \{a_{11}, a_{12}, a_1\}$, $g = \{a_4, a_{13}, a_{10}\}$. They have a (quantum) realization in \mathbb{R}^3 consisting of the 13 projections associated with the one dimensional subspaces spanned by the vectors from the origin $(0, 0, 0)^T$ to $a_1 = (1, \sqrt{2}, 0)^T$, $a_2 = (\sqrt{2}, -1, -3)^T$, $a_3 = (\sqrt{2}, -1, 1)^T$, $a_4 = (0, 1, 1)^T$, $a_5 = (\sqrt{2}, 1, -1)^T$, $a_6 = (\sqrt{2}, 1, 3)^T$, $a_7 = (-1, \sqrt{2}, 0)^T$, $a_8 = (\sqrt{2}, 1, -3)^T$, $a_9 = (\sqrt{2}, 1, 1)^T$, $a_{10} = (0, 1, -1)^T$, $a_{11} = (\sqrt{2}, -1, -1)^T$, $a_{12} = (\sqrt{2}, -1, 3)^T$, $a_{13} = (1, 0, 0)^T$, respectively [533, p. 206, Fig. 1] (see also [523, Fig. 4, p. 5387])

Table 12.4 The 14 two-valued states on the Specker bug (cat’s cradle) logic

#	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}
v_1	1	0	0	0	1	0	0	0	1	0	0	0	1
v_2	1	0	0	1	0	1	0	0	1	0	0	0	0
v_3	1	0	0	0	1	0	0	1	0	1	0	0	0
v_4	0	1	0	0	1	0	0	0	1	0	0	1	1
v_5	0	1	0	0	1	0	0	1	0	0	1	0	1
v_6	0	1	0	1	0	1	0	0	1	0	0	1	0
v_7	0	1	0	1	0	0	1	0	0	0	1	0	0
v_8	0	1	0	1	0	1	0	1	0	0	1	0	0
v_9	0	1	0	0	1	0	0	1	0	1	0	1	0
v_{10}	0	0	1	0	0	0	1	0	0	0	1	0	1
v_{11}	0	0	1	0	0	1	0	1	0	0	1	0	1
v_{12}	0	0	1	0	0	1	0	0	1	0	0	1	1
v_{13}	0	0	1	0	0	0	1	0	0	1	0	1	0
v_{14}	0	0	1	0	0	1	0	1	0	1	0	1	0

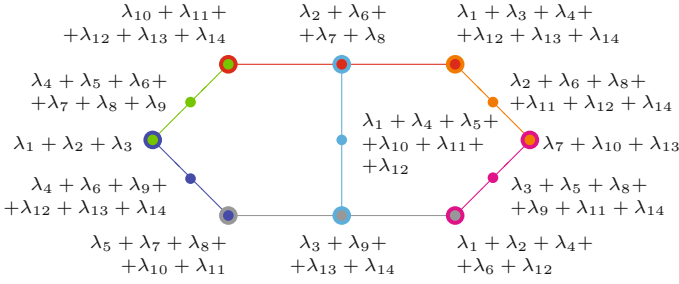


Fig. 12.10 Classical probabilities on the Specker bug (cat’s cradle) logic; $\lambda_1 + \dots + \lambda_{14} = 1$, $0 \leq \lambda_1, \dots, \lambda_{14} \leq 1$, taken from Ref. [521]. The two-valued states $i = 1, \dots, 14$ can be identified by taking $\lambda_j = \delta_{i,j}$ for all $j = 1, \dots, 14$

$$p_1 + p_7 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_7 + \lambda_{10} + \lambda_{13} \leq \sum_{i=1}^{14} \lambda_i = 1. \tag{12.42}$$

For two-valued measures this yields the “1–0” or “true implies false” rule [515]: if a_1 is true, then a_7 must be false. For the sake of another proof by contradiction, suppose a_1 as well as a_7 were both true. This would (by the admissibility rules) imply a_3, a_5, a_9, a_{11} to be false, which in turn would imply both a_4 as well as a_{10} , which have to be true in one and the same context – a clear violation of the admissibility rules stating that within a single context there can only be atom which is true. This property, which has already been exploited by Kochen and Specker [314, Γ_1] to construct both a logic with a non-separating, as well as one with a non-existent set of two valued states. These former case will be discussed in the next section. For the time being, instead of drawing all two valued states separately, Fig. 12.10 enumerates the classical probabilities on the Specker bug (cat’s cradle) logic.

The hull problem yields 23 facet inequalities; one of them relating p_1 to p_7 : $p_1 + p_2 + p_7 + p_6 \geq 1 + p_4$, which is satisfied, since, by subadditivity, $p_1 + p_2 = 1 - p_3$, $p_7 + p_6 = 1 - p_5$, and $p_4 = 1 - p_5 - p_3$. This is a good example of a situation in which considering just Boole–Bell type inequalities do not immediately reveal important aspects of the classical probabilities on such logics.

A restricted hull calculation for the joint expectations on the six edges of the Greechie orthogonality diagram yields 18 inequalities; among them

$$E_{13} + E_{57} + E_{9,11} \leq E_{35} + E_{79} + E_{11,1}. \tag{12.43}$$

A tightened “true implies 3-times-false” logic depicted in Fig. 12.11 has been introduced by Yu and Oh [584]. As can be derived from admissibility in a straightforward manner, the set of 24 two-valued states [536] enforces at most one of the four atoms h_0, h_1, h_2, h_3 to be 1. Therefore, classically $p_{h_0} + p_{h_1} + p_{h_2} + p_{h_3} \leq 1$. This can also be explicitly demonstrated by noticing that, from the 24 two-valued states, exactly 3 acquire the value 1 on each one of the four atoms h_0, h_1, h_2 , and h_3 ;

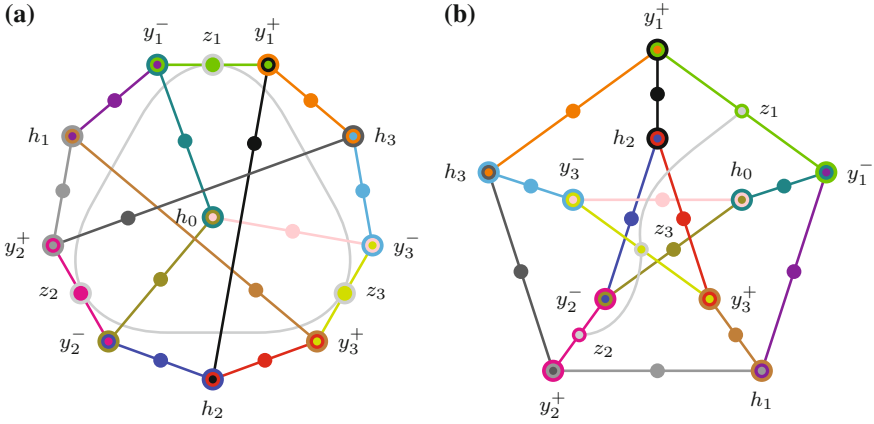


Fig. 12.11 Two equivalent representations of a Petersen graph-like (with one additional context connecting $z_1, z_2,$ and z_3) Greechie diagram of the logic considered by Yu and Oh [584, Fig. 2]. The set of two-valued states enforces at most one of the four atoms h_0, h_1, h_2, h_3 to be 1. The logic has a (quantum) realization in \mathbb{R}^3 consisting of the 25 projections; associated with the one dimensional subspaces spanned by the 13 vectors from the origin $(0, 0, 0)^T$ to $z_1 = (1, 0, 0)^T, z_2 = (0, 1, 0)^T, z_3 = (0, 0, 1)^T, y_1^- = (0, 1, -1)^T, y_2^- = (1, 0, -1)^T, y_3^- = (1, -1, 0)^T, y_1^+ = (0, 1, 1)^T, y_2^+ = (1, 0, 1)^T, y_3^+ = (1, 1, 0)^T, h_0 = (1, 1, 1)^T, h_1 = (-1, 1, 1)^T, h_2 = (1, -1, 1)^T, h_3 = (1, 1, -1)^T$, respectively [584]

also the respective two-valued states are different for these four different atoms $h_0, h_1, h_2,$ and h_3 . More explicitly, suppose the set of two-valued states is enumerated in such a way that the respective probabilities on the atoms $h_0, h_1, h_2,$ and h_3 are $p_{h_0} = \lambda_1 + \lambda_2 + \lambda_3, p_{h_1} = \lambda_4 + \lambda_5 + \lambda_6, p_{h_2} = \lambda_7 + \lambda_8 + \lambda_9,$ and $p_{h_3} = \lambda_{10} + \lambda_{11} + \lambda_{12}$. Because of additivity the 24 positive weights $\lambda_1, \dots, \lambda_{24} \geq 0$ must add up to 1; that is, $\sum_{i=1}^{24} \lambda_i = 1$. Therefore [compare with Eq. (12.42)],

$$p_{h_0} + p_{h_1} + p_{h_2} + p_{h_3} = \sum_{j=1}^{12} \lambda_j \leq \sum_{i=1}^{24} \lambda_i = 1. \tag{12.44}$$

Tkadlec has noted [536] that Fig. 12.11 contains 3 Specker bug subdiagrams per atom h_i , thereby rendering the “true implies 3-times-false” property. For instance, for h_1 the three Specker bugs are formed by the three sets of contexts (missing non-interwining atoms should be added)

- 1 : $\{\{h_1, y_3^+\}, \{y_3^+, y_3^-\}, \{y_3^-, h_3\}, \{h_3, y_1^+\}, \{y_1^+, y_1^-\}, \{y_1^-, h_1\}, \{z_1, z_3\}\},$
- 2 : $\{\{h_1, y_2^+\}, \{y_2^+, y_2^-\}, \{y_2^-, h_2\}, \{h_2, y_1^+\}, \{y_1^+, y_1^-\}, \{y_1^-, h_1\}, \{z_1, z_2\}\},$ (12.45)
- 3 : $\{\{h_1, y_3^+\}, \{y_3^+, y_3^-\}, \{y_3^-, h_0\}, \{h_0, y_2^-\}, \{y_2^-, y_2^+\}, \{y_2^+, h_1\}, \{z_3, z_2\}\}.$

Other “true implies true” logics have been introduced by Belinfante [39, Fig. C.1. p. 67], Pitowsky [289, p. 394], Clifton [1, 293, 546], as well as Cabello and G. García-Alcaine [100, Lemma 1].

Notice that, if a second Specker bug logic is placed along b_1 and b_7 , just as in the Kochen–Specker Γ_3 logic [314, p. 70], this imposes an additional “true implies false” condition; together with the “true implies false” condition of the first logic this implies the fact that a_1 and a_7 can no longer be separated by some two-valued state: whenever one is true, the other one must be true as well, and *vice versa*. This Kochen–Specker logic Γ_3 will be discussed in the next Sect. 12.9.8.6.

Notice further that if we manage to iterate this process in such a manner that, with every i th iteration we place another Kochen–Specker Γ_3 logic along b_i , while at the same time increasing the angle between b_i and b_1 , then eventually we shall arrive at a situation in which b_1 and b_i are part of a context (in terms of Hilbert space: they correspond to orthogonal vectors). But admissibility disallows two-valued measures with more than one, and in particular, two “true” atoms within a single block. As a consequence, if such a configuration is realizable (say, in 3-dimensional Hilbert space), then it cannot have any two-valued state satisfying the admissibility criteria. This is the Kochen–Specker theorem, as exposed in the Kochen–Specker Γ_3 logic [314, p. 69], which will be discussed in Sect. 12.9.8.7.

12.9.8.6 Combo of Two Linked Specker Bug Logics Inducing Non-separability

As we are heading toward logics with less and less “rich” set of two-valued states we are approaching a logic depicted in Fig. 12.13 which is a combination of two Specker bug logics linked by two external contexts. It is the Γ_3 -configuration of Kochen–Specker [314, p. 70] with a set of two-valued states which is no longer separating: In this case one obtains the “one-one” and “zero-zero rules” [515], stating that a_1 occurs if and only if b_1 occurs (likewise, a_7 occurs if and only if b_7 occurs): Suppose v is a two-valued state on the Γ_3 -configuration of Kochen–Specker. Whenever $v(a_1) = 1$, then $v(c) = 0$ because it is in the same context $\{a_1, c, b_7\}$ as a_1 . Furthermore, because of Eq. (12.41), whenever $v(a_1) = 1$, then $v(a_7) = 0$. Because b_1 is in the same context $\{a_7, c, b_1\}$ as a_7 and c , because of admissibility, $v(b_1) = 1$. Conversely, by symmetry, whenever $v(b_1) = 1$, so must be $v(a_1) = 1$. Therefore it can never happen that either one of the two atoms a_1 and b_1 have different dichotomic values. (Eq. 12.46 is compatible with these value assignments.) The same is true for the pair of atoms a_7 and b_7 .

Note that one needs two Specker bug logics tied together (at their “true implies false” extremities) to obtain non-separability; just extending one to the Kochen–Specker Γ_1 logic [314, p. 68] of Fig. 12.12 discussed earlier to obtain “true implies true” would be insufficient. Because in this case a consistent two-valued state exists for which $v(b_1) = v(b_7) = 1$ and $v(a_1) = v(a_7) = 0$, thereby separating a_1 from b_1 , and *vice versa*. A second Specker bug logic is needed to eliminate this case; in particular, $v(b_1) = v(b_7) = 1$.

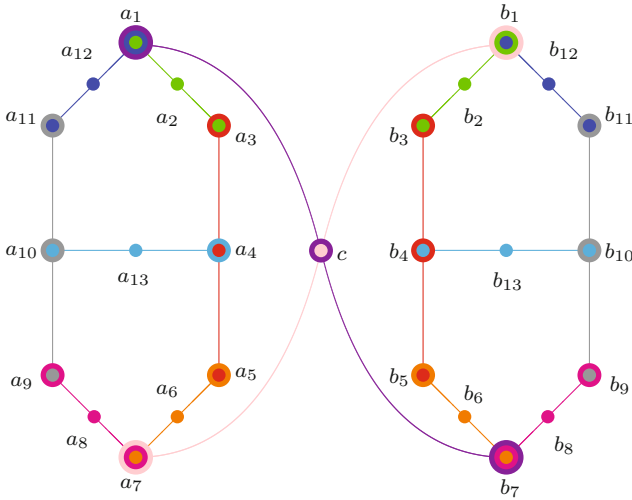


Fig. 12.13 Greechie diagram of two linked Specker bug (cat’s cradle) logics G_3 . The logic has a (quantum) realization in \mathbb{R}^3 consisting of the 27 projections associated with the one dimensional subspaces spanned by the vectors from the origin $(0, 0, 0)^\top$ to the 13 points mentioned in Fig. 12.9, the 3 points mentioned in Fig. 12.12, as well as $b_2 = (1, -\sqrt{2}, -3)^\top$, $b_3 = (-1, \sqrt{2}, -1)^\top$, $b_4 = (1, 0, -1)^\top$, $b_5 = (1, \sqrt{2}, 1)^\top$, $b_6 = (1, \sqrt{2}, -3)^\top$, $b_8 = (1, \sqrt{2}, 3)^\top$, $b_9 = (1, \sqrt{2}, -1)^\top$, $b_{10} = (1, 0, 1)^\top$, $b_{11} = (-1, \sqrt{2}, 1)^\top$, $b_{12} = (-1, \sqrt{2}, -3)^\top$, $b_{13} = (0, 1, 0)^\top$, respectively [533, p. 206, Fig. 1]. Note that, with this realization, there is an additional context $\{a_{13}, c, b_{13}\}$ not drawn here, which imposes an additional constraint $v(a_{13}) + v(c) + v(b_{13}) = 1$ on any two-valued measure v (See also the proof of Proposition 7.2 in Ref. [523, p. 5392].)

Besides the quantum mechanical realization of this logic in terms of propositions which are projection operators corresponding to vectors in three-dimensional Hilbert space suggested by Kochen and Specker [314], Tkadlec has given [533, p. 206, Fig. 1] an explicit collection of such vectors (see also the proof of Proposition 7.2 in Ref. [523, p. 5392]).

Probabilistic Criteria Against Value Definiteness from Constraints on Two-Valued measures

The “1-1” or “true implies true” rule can be taken as an operational criterion for quantization: Suppose that one prepares a system to be in a pure state corresponding to a_1 , such that the preparation ensures that $v(a_1) = 1$. If the system is then measured along b_1 , and the proposition that the system is in state b_1 is found to be *not* true, meaning that $v(b_1) \neq 1$ (the respective detector does not click), then one has established that the system is not performing classically, because classically the set of two-valued states requires non-separability; that is, $v(a_1) = v(b_1) = 1$. With the Tkadlec directions taken from Figs. 12.9 and 12.12, $|a_1\rangle = (1/\sqrt{3}) (1, \sqrt{2}, 0)^\top$

and $|\mathbf{b}_1\rangle = (1/\sqrt{3}) (\sqrt{2}, 1, 0)^T$ so that the probability to find a quantized system prepared along $|\mathbf{a}_1\rangle$ and measured along $|\mathbf{b}_1\rangle$ is $p_{a_1}(b_1) = |\langle \mathbf{b}_1 | \mathbf{a}_1 \rangle|^2 = 8/9$, and that a violation of classicality should occur with probability $1/9$. Of course, any other classical prediction, such as the “1-0” or “true implies false” rule, or more general classical predictions such as of Eq. (12.41) can also be taken as empirical criteria for non-classicality [521, Sect. 11.3.2.].

Indeed, already Stairs [483, pp. 588–589] has argued along similar lines for the Specker bug “true implies false” logic (a translation into our nomenclature is: $m1(1) \equiv a_1, m2(1) \equiv a_3, m2(2) \equiv a_5, m2(3) \equiv a_4, m3(1) \equiv a_{11}, m3(2) \equiv a_9, m3(3) \equiv a_{10}, m4(1) \equiv a_7$). Independently Clifton (there is a note added in proof to Stairs [483, pp. 588–589]) presents a similar argument, based upon (i) another “true implies true” logic [1, 293, 546, Sects. II, III, Fig. 1] inspired by Bell [39, Fig. C.1 p. 67] (cf. also Pitowsky [289, p. 394]), as well as (ii) on the Specker bug logic [1, Sect. IV, Fig. 2]. More recently Hardy [70, 264, 265] as well as Cabello and García-Alcaine and others [24, 90, 95, 96, 99, 138] discussed such scenarios. These criteria for non-classicality are benchmarks aside from the Boole–Bell type polytope method, and also different from the full Kochen–Specker theorem.

Imbedability

As every algebra imbeddable in a Boolean algebra must have a separating set of two valued states, this logic is no longer “classical” in the sense of “homomorphically (structure-preserving) imbeddable.” Nevertheless, two-valued states can still exist. It is just that these states can no longer differentiate between the pairs of atoms (a_1, b_1) as well as (a_7, b_7) . Partition logics and their generalized urn or finite automata models fail to reproduce two linked Specker bug logics resulting in a Kochen–Specker Γ_3 logic even at this stage. Of course, the situation will become more dramatic with the non-existence of any kind of two-valued state (interpretable as truth assignment) on certain logics associate with quantum propositions.

Complementarity and non-distributivity is not enough to characterize logics which do not have a quasi-classical (partition logical, set theoretical) interpretation. While in a certain, graph coloring sense the “richness/scarcity” and the “number” of two-valued homomorphisms” yields insights into the old problem of the structural property [152] by separating quasi-classical from quantum logics, the problem of finding smaller, maybe minimal, subsets of graphs with a non-separating set of two-valued states still remains an open challenge.

Chromatic Inseparability

The “true implies true” rule is associated with *chromatic separability*; in particular, with the impossibility to separate two atoms a_7 and b_7 with less than four colors. A proof is presented in Fig. 12.14. That chromatic separability on the unit sphere requires 4 colors is implicit in Refs. [245, 269].

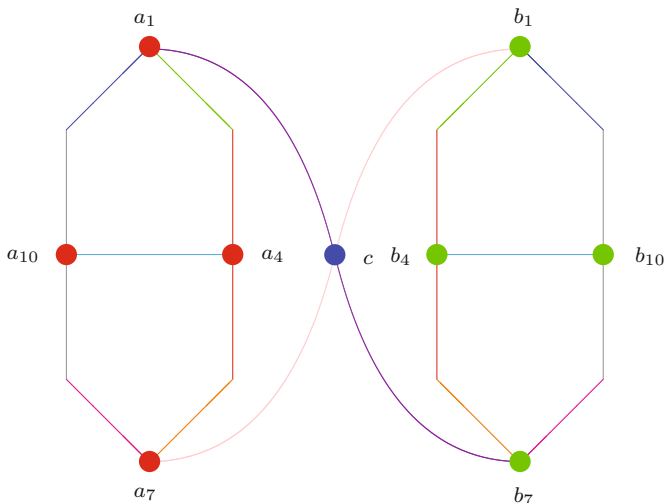


Fig. 12.14 Proof (by contradiction) that chromatic separability of two linked Specker bug (cat’s cradle) logics Γ_3 cannot be achieved with three colors. In particular, a_7 and b_7 cannot be separated, as this would result in the depicted inconsistent coloring: suppose a red/green/blue coloring with chromatic admissibility (“all three colors occur only once per context or block or Boolean sub-algebra”) is possible. Then, if a_7 is colored red and b_7 is colored green, c must be colored blue. Therefore, a_1 must be colored red. Therefore, a_4 as well as a_{10} must be colored red (similar for green on the second Specker bug), contradicting admissibility

12.9.8.7 Propositional Structures Without Two-Valued States

Gleason-Type Continuity

Gleason’s theorem [240] was a response to Mackey’s problem to “determine all measures on the closed subspaces of a Hilbert space” contained in a review [351] of Birkhoff and von Neumann’s centennial paper [62] on the logic of quantum mechanics. Starting from von Neumann’s formalization of quantum mechanics [552, 554], the quantum mechanical probabilities and expectations (aka the Born rule) are essentially derived from (sub)additivity among the quantum context; that is, from sub-classicality: within any context (Boolean subalgebra, block, maximal observable, orthonormal base) the quantum probabilities sum up to 1.

Gleason’s finding caused ripples in the community, at least of those who cared and coped with it [41, 151, 180, 301, 314, 401, 434, 591]. (I recall having an argument with Van Lambalgen around 1983, who could not believe that anyone in the larger quantum community had not heard of Gleason’s theorem. As we approached an elevator at Vienna University of Technology’s Freihaus building we realized there was also one very prominent member of Vienna experimental community entering the cabin. I suggested to stage an example by asking; and *voila*. . .)

With the possible exception of Specker who did not explicitly refer to the Gleason's theorem in independently announcing that two-valued states on quantum logics cannot exist [479] – he must have made up his mind from other arguments and preferred to discuss scholastic philosophy; at that time the Swiss may have had their own biotope – Gleason's theorem directly implies the absence of two-valued states. Indeed, at least for finite dimensions [11, 12], as Zierler and Schlessinger [591, p. 259, Example 3.2] (even before publication of Bell's review [41]) noted, “*it should also be mentioned that, in fact, the non-existence of two-valued states is an elementary geometric fact contained quite explicitly in [240, Paragraph 2.8].*”

Now, Gleason's Paragraph 2.8 contains the following main (necessity) theorem [240, p. 888]: “*Every non-negative frame function on the unit sphere S in \mathbb{R}^3 is regular.*” Whereby [240, p. 886] “*a frame function f [[satisfying additivity]] is regular if and only if there exists a self-adjoint operator \mathbf{T} defined on [[the separable Hilbert space]] \mathfrak{H} such that $f(|x\rangle) = \langle \mathbf{T}x|x\rangle$ for all unit vectors $|x\rangle$.*” (Of course, Gleason did not use the Dirac notation.)

In what follows we shall consider Hilbert spaces of dimension $n = 3$ and higher. Suppose that the quantum system is prepared to be in a pure state associated with the unit vector $|x\rangle$, or the projection operator $|x\rangle\langle x|$.

As all self-adjoint operators have a spectral decomposition [260, Sect. 79], and the scalar product is (anti)linear in its arguments, let us, instead of \mathbf{T} , only consider one-dimensional orthogonal projection operators $\mathbf{E}_i^2 = \mathbf{E}_i = |y_i\rangle\langle y_i|$ (formed by the unit vector $|y_i\rangle$ which are elements of an orthonormal basis $\{|y_1\rangle, \dots, |y_n\rangle\}$) occurring in the spectral sum of $\mathbf{T} = \sum_{i=1}^{n \geq 3} \lambda_i \mathbf{E}_i$, with $\mathbb{I}_n = \sum_{i=1}^{n \geq 3} \mathbf{E}_i$.

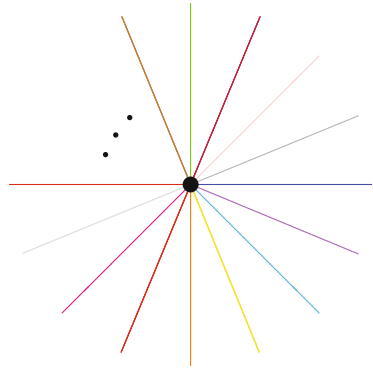
Thus if \mathbf{T} is restricted to some one-dimensional projection operator $\mathbf{E} = |y\rangle\langle y|$ along $|y\rangle$, then Gleason's main theorem states that any frame function reduces to the absolute square of the scalar product; and in real Hilbert space to the square of the angle between those vectors spanning the linear subspaces corresponding to the two projectors involved; that is (note that \mathbf{E} is self-adjoint), $f_y(|x\rangle) = \langle \mathbf{E}x|x\rangle = \langle x|\mathbf{E}x\rangle = \langle x|y\rangle\langle y|x\rangle = |\langle x|y\rangle|^2 = \cos^2 \angle(x, y)$.

Hence, unless a configuration of contexts is not of the star-shaped Greechie orthogonality diagram form – meaning that they all share one common atom; and, in terms of geometry, meaning that all orthonormal bases share a common vector – and the two-valued state has value 1 on its centre, as depicted in Fig. 12.15, there is no way how any two contexts could have a two-valued assignment; even if one context has one: it is just not possible by the continuous, \cos^2 -form of the quantum probabilities. That is (at least in this author's believe) the watered down version of the remark of Zierler and Schlessinger [591, p. 259, Example 3.2].

Finite Logics Admitting No Two-Valued States

When it comes to the absence of a global two-valued state on quantum logics corresponding to Hilbert spaces of dimension three and higher – where contexts or blocks can be intertwined or pasted [376] to form chains – Kochen and Specker [314] pursued a very concrete, “constructive” (in the sense of finitary mathematical objects

Fig. 12.15 Greechie diagram of a star shaped configuration with a variety of contexts, all intertwined in a single “central” atom; with overlaid two-valued state (bold black filled circle) which is one on the centre atom and zero everywhere else (see also Refs. [3, 5, 6])



but not in the sense of physical operationalizability [79]) strategy: they presented finite logics realizable by vectors (from the origin to the unit sphere) spanning one-dimensional subspaces, equivalent to observable propositions, which allowed for lesser and lesser two-valued state properties. For the reason of non-imbedability is already enough to consider two linked Specker bugs logics Γ_3 [314, p. 70], as discussed in Sect. 12.9.8.6.

Kochen and Specker went further and presented a proof by contradiction of the non-existence of two-valued states on a finite number of propositions, based on their Γ_1 “true implies true” logic [314, p. 68] discussed in Fig. 12.12, iterating them until they reached a complete contradiction in their Γ_2 logic [314, p. 69]. As has been pointed out earlier, their representation as points of the sphere is a little bit “buggy” (as could be expected from the formation of so many bugs): as Tkadlec has observed, Kochen–Specker diagram Γ_2 it is not a one-to-one representation of the logic, because some different points at the diagram represent the same element of corresponding orthomodular poset (cf. Ref. [523, p. 5390], and Ref. [535, p.]).

The early 1990s saw an ongoing flurry of papers recasting the Kochen–Specker proof with ever smaller numbers of, or more symmetric, configurations of observables (see Refs. [17, 83, 96, 97, 112, 307, 340, 364, 385, 386, 390, 391, 408, 472, 523, 533–535, 557, 558, 583, 593] for an incomplete list). Arguably the most compact such logic is one in four-dimensional space suggested by Cabello, Estebaranz and García-Alcaine [91, 96, 385]. It consists of 9 contexts, with each of the 18 atoms tightly intertwined in two contexts. Its Greechie orthogonality diagram is drawn in Fig. 12.16.

In a parity proof by contradiction consider the particular subset of real four-dimensional Hilbert space with a “parity property,” consisting of 18 atoms a_1, \dots, a_{18} in 9 contexts, as depicted in Fig. 12.16. Note that, on the one hand, each atom/point/vector/projector belongs to exactly two – that is, an *even* number of – contexts; that is, it is biconnected. Therefore, any enumeration of all the contexts occurring in the graph depicted in Fig. 12.16 would contain an *even* number of 1s assigned. Because, due to non-contextuality and biconnectivity, any atom a with $v(a) = 1$ along one

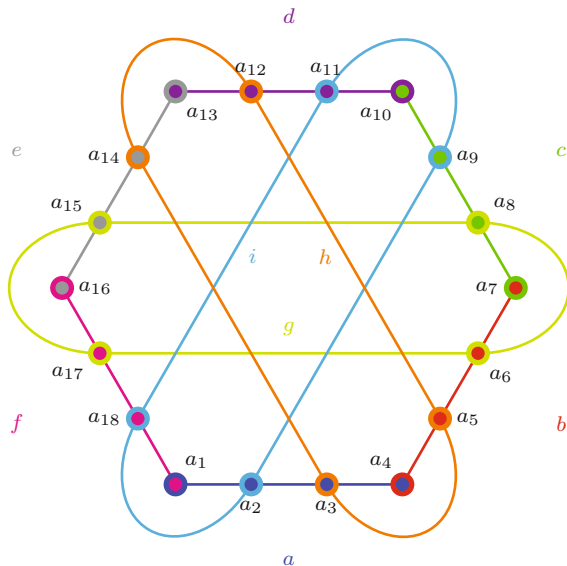


Fig. 12.16 The most compact way of deriving the Kochen–Specker theorem in four dimensions has been given by Cabello, Estebaranz and García-Alcaine [96]. The configuration consists of 18 biconnected (two contexts intertwine per atom) atoms a_1, \dots, a_{18} in 9 contexts. It has a (quantum) realization in \mathbb{R}^4 consisting of the 18 projections associated with the one dimensional subspaces spanned by the vectors from the origin $(0, 0, 0, 0)^T$ to $a_1 = (0, 0, 1, -1)^T$, $a_2 = (1, -1, 0, 0)^T$, $a_3 = (1, 1, -1, -1)^T$, $a_4 = (1, 1, 1, 1)^T$, $a_5 = (1, -1, 1, -1)^T$, $a_6 = (1, 0, -1, 0)^T$, $a_7 = (0, 1, 0, -1)^T$, $a_8 = (1, 0, 1, 0)^T$, $a_9 = (1, 1, -1, 1)^T$, $a_{10} = (-1, 1, 1, 1)^T$, $a_{11} = (1, 1, 1, -1)^T$, $a_{12} = (1, 0, 0, 1)^T$, $a_{13} = (0, 1, -1, 0)^T$, $a_{14} = (0, 1, 1, 0)^T$, $a_{15} = (0, 0, 0, 1)^T$, $a_{16} = (1, 0, 0, 0)^T$, $a_{17} = (0, 1, 0, 0)^T$, $a_{18} = (0, 0, 1, 1)^T$, respectively [92, Fig. 1] (for alternative realizations see Refs. [91, 92])

context must have the same value 1 along the second context which is intertwined with the first one – to the values 1 appear in pairs.

Alas, on the other hand, in such an enumeration there are nine – that is, an *odd* number of – contexts. Hence, in order to obey the quantum predictions, any two-valued state (interpretable as truth assignment) would need to have an *odd* number of 1s – exactly one for each context. Therefore, there cannot exist any two-valued state on Kochen–Specker type graphs with the “parity property.”

More concretely, note that, within each one of those 9 contexts, the sum of any state on the atoms of that context must add up to 1. That is, due to additivity (12.24) and (12.25) one obtains a system of 9 equations

$$\begin{aligned}
v(a) &= v(a_1) + v(a_2) + v(a_3) + v(a_4) = 1, \\
v(b) &= v(a_4) + v(a_5) + v(a_6) + v(a_7) = 1, \\
v(c) &= v(a_7) + v(a_8) + v(a_9) + v(a_{10}) = 1, \\
v(d) &= v(a_{10}) + v(a_{11}) + v(a_{12}) + v(a_{13}) = 1, \\
v(e) &= v(a_{13}) + v(a_{14}) + v(a_{15}) + v(a_{16}) = 1, \\
v(f) &= v(a_{16}) + v(a_{17}) + v(a_{18}) + v(a_1) = 1, \\
v(g) &= v(a_6) + v(a_8) + v(a_{15}) + v(a_{17}) = 1, \\
v(h) &= v(a_3) + v(a_5) + v(a_{12}) + v(a_{14}) = 1, \\
v(i) &= v(a_2) + v(a_9) + v(a_{11}) + v(a_{18}) = 1.
\end{aligned} \tag{12.47}$$

By summing up the left hand side and the right hand sides of the equations, and since all atoms are biconnected, one obtains

$$2 \left[\sum_{i=1}^{18} v(a_i) \right] = 9. \tag{12.48}$$

Because $v(a_i) \in \{0, 1\}$ the sum in (12.48) must add up to some natural number M . Therefore, Eq. (12.48) is impossible to solve in the domain of natural numbers, as on the left and right hand sides there appear even ($2M$) and odd (9) numbers, respectively.

Of course, one could also prove the nonexistence of any two-valued state (interpretable as truth assignment) by exhaustive attempts (possibly exploiting symmetries) to assign values 0s and 1s to the atoms/points/vectors/projectors occurring in the graph in such a way that both the quantum predictions as well as context independence is satisfied. This latter method needs to be applied in cases with Kochen–Specker type diagrams without the “parity property;” such as in the original Kochen–Specker proof [314]. (However, admissibility (IV) is too weak for a proof of this type, as it allows also a third, value indefinite, state, which spoils the arguments [6].)

This result, as well as the original Kochen–Specker theorem, is *state independent* insofar as it applies to an arbitrary quantum state. One could *reduce the size of the proof* by assuming a particular state. Such proofs are called *state-specific* or *state dependent*. By following Cabello, Estebarez and García-Alcaine [96, Eqs. (10)–(19), p. 185] their state independent proof utilizing the logic depicted in Fig. 12.16 can be transferred to a state-specific proof as follows: suppose that the quantum (or quanta, depending upon the physical realization) is prepared in the state

$$v(a_1) = 1, \tag{12.49}$$

so that any two-valued state must obey the admissibility rules

$$v(a_2) = v(a_3) = v(a_4) = v(a_{16}) = v(a_{17}) = v(a_{18}) = 0. \tag{12.50}$$

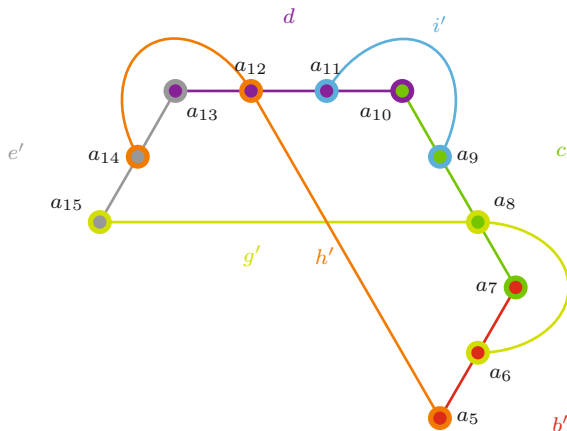


Fig. 12.17 Greechie orthogonality diagram of a state-specific proof of the Kochen–Specker theorem based on the assumption that the physical system is in state a_1 , such that $v(a_1) = 1$. The additivity and admissibility constraints (12.51) represent different “reduced” (or “truncated”) contexts, because all states $v(a_2) = v(a_3) = v(a_4) = v(a_{16}) = v(a_{17}) = v(a_{18}) = 0$ “orthogonal to” a_1 must vanish

The additivity relations (12.47) reduce to seven equations (two equations encoding contexts a and f are satisfied trivially)

$$\begin{aligned}
 v(b') &= v(a_5) + v(a_6) + v(a_7) = 1, \\
 v(c) &= v(a_7) + v(a_8) + v(a_9) + v(a_{10}) = 1, \\
 v(d) &= v(a_{10}) + v(a_{11}) + v(a_{12}) + v(a_{13}) = 1, \\
 v(e') &= v(a_{13}) + v(a_{14}) + v(a_{15}) = 1, \\
 v(g') &= v(a_6) + v(a_8) + v(a_{15}) = 1, \\
 v(h') &= v(a_5) + v(a_{12}) + v(a_{14}) = 1, \\
 v(i') &= v(a_9) + v(a_{11}) = 1.
 \end{aligned}
 \tag{12.51}$$

The configuration is depicted in Fig. 12.17. As all atoms remain to be biconnected and there are 7, that is, an odd number, of equations, value indefiniteness can be proven by a similar parity argument as before. One could argue that the “primed” contexts in (12.51) are not complete because those contexts are “truncated.” However, every completion would result in vectors orthogonal to a_1 ; and therefore their values must again be zero.

Chromatic Number of the Sphere

Graph coloring allows another view on value (in)definiteness. The *chromatic number* of a graph is defined as the least number of colors needed in any total coloring of a graph; with the constraint that two adjacent vertices have distinct colors.

Suppose that we are interested in the chromatic number of graphs associated with both (i) the real and (ii) the rational three-dimensional unit sphere.

More generally, we can consider n -dimensional unit spheres with the same adjacency property defined by orthogonality. An orthonormal basis will be called context (block, maximal observable, Boolean subalgebra), or, in this particular area, a n -clique. Note that for any such graphs involving n -cliques the chromatic number of this graph is at least be n (because the chromatic number of a single n -clique or context is n).

Thereby vertices of the graph are identified with points on the three-dimensional unit sphere; with adjacency defined by orthogonality; that is, two vertices of the graph are adjacent if and only if the unit vectors from the origin to the respective two points are orthogonal.

The connection to quantum logic is this: any context (block, maximal observable, Boolean subalgebra, orthonormal basis) can be represented by a triple of points on the sphere such that any two unit vectors from the origin to two distinct points of that triple of points are orthogonal. Thus graph adjacency in logical terms indicates “belonging to some common context (block, maximal observable, Boolean subalgebra, orthonormal basis).”

In three dimensions, if the chromatic number of graphs is four or higher, there does not globally exist any consistent coloring obeying the rule that adjacent vertices (orthogonal vectors) must have different colors: if one allows only three different colors, then somewhere in that graph of chromatic number higher than three, adjacent vertices must have the same colors (or else the chromatic number would be three or lower).

By a similar argument, non-separability of two-valued states – such as encountered in Sect. 12.9.8.6 with the Γ_3 -configuration of Kochen–Specker [314, p. 70] – translates into non-differentiability by colorings with colors less or equal to the number of atoms in a block (cf. Fig. 12.14).

Godsil and Zaks [245, 269] proved the following results:

1. the chromatic number of the graph based on points of real-valued unit sphere is four [245, Lemma 1.1].
2. the chromatic number of rational points on the unit sphere $S^3 \cap \mathbb{Q}^3$ is three [245, Lemma 1.2].

We shall concentrate on (i) and discuss (ii) later. As has been pointed out by Godsil in an email conversation from March 13, 2016 [244], “*the fact that the chromatic number of the unit sphere in \mathbb{R}^3 is four is a consequence of Gleason’s theorem, from which the Kochen–Specker theorem follows by compactness. Gleason’s result implies that there is no subset of the sphere that contains exactly one point from each orthonormal basis.*”

Indeed, any coloring can be mapped onto a two-valued state by identifying a single color with “1” and all other colors with “0.” By reduction, all propositions on two-valued states translate into statements about graph coloring. In particular, if the chromatic number of any logical structure representable as graph consisting of

n -atomic contexts (blocks, maximal observables with n outcomes, Boolean subalgebras 2^n , orthonormal bases with n elements) – for instance, as Greechie orthogonality diagram of quantum logics – is larger than n , then there cannot be any globally consistent two-valued state (truth value assignment) obeying adjacency (aka admissibility). Likewise, if no two-valued states on a logic which is a pasting of n -atomic contexts exist, then, by reduction, no global consistent coloring with n different colors exists. Therefore, the Kochen–Specker theorem proves that the chromatic number of the graph corresponding to the unit sphere with adjacency defined as orthogonality must be higher than three.

Based on Godsil and Zaks finding that the chromatic number of rational points on the unit sphere $S^3 \cap \mathbb{Q}^3$ is three [245, Lemma 1.2] – thereby constructing a two-valued measure on the rational unit sphere by identifying one color with “1” and the two remaining colors with “0” – there exist “exotic” options to circumvent Kochen–Specker type constructions which have been quite aggressively (Cabello has referred to this as the *second contextuality war* [94]) marketed by allegedly “nullifying” [369] the respective theorems under the umbrella of “finite precision measurements” [32, 75, 76, 146, 306, 366]: the support of vectors spanning the one-dimensional subspaces associated with atomic propositions could be “diluted” yet dense, so much so that the intertwines of contexts (blocks, maximal observables, Boolean subalgebras, orthonormal bases) break up; and the contexts themselves become “free and isolated.” Under such circumstances the logics decay into horizontal sums; and the Greechie orthogonality diagrams are just disconnected stacks of previously intertwined contexts. As can be expected, proofs of Gleason- or Kochen–Specker-type theorems do no longer exist, as the necessary intertwines are missing.

The “nullification” claim and subsequent ones triggered a lot of papers, some cited in [32]; mostly critical – of course, not to the results of Godsil and Zaks’s finding (ii); how could they? – but to their physical applicability. Peres even wrote a parody by arguing that “*finite precision measurement nullifies Euclid’s postulates*” [392], so that “nullification” of the Kochen–Specker theorem might have to be our least concern.

Exploring Value Indefiniteness

Maybe one could, with all due respect, speak of “extensions” of the Kochen–Specker theorem by looking at situations in which a system is prepared in a state $|\mathbf{x}\rangle\langle\mathbf{x}|$ along direction $|\mathbf{x}\rangle$ and measured along a non-orthogonal, non-collinear projection $|\mathbf{y}\rangle\langle\mathbf{y}|$ along direction $|\mathbf{y}\rangle$. Those extensions yield what may be called [286, 401] *indeterminacy*. Indeterminacy may be just another word for *contextuality*; but, as has been suggested by the realist Bell, the latter term implicitly implies that there “is something (rather than nothing) out there,” some “pre-existing observable” which, however, needs to depend on the context of the measurement. To avoid such implicit assumption we shall henceforth use *indeterminacy* rather than *contextuality*.

Pitowsky’s *logical indeterminacy principle* [401, Theorem 6, p. 226] states that, given two linearly independent non-orthogonal unit vectors $|\mathbf{x}\rangle$ and $|\mathbf{y}\rangle$ in \mathbb{R}^3 , there is a finite set of unit vectors $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ containing $|\mathbf{x}\rangle$ and $|\mathbf{y}\rangle$ for which the following statements hold:

1. There is no (not necessarily two-valued) state v on $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ which satisfies $v(|\mathbf{x}\rangle) = v(|\mathbf{y}\rangle) = 1$.
2. There is no (not necessarily two-valued) v on $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ which satisfies $v(|\mathbf{x}\rangle) = 1$ and $v(|\mathbf{y}\rangle) = 0$.
3. There is no (not necessarily two-valued) state v on $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ which satisfies $v(|\mathbf{x}\rangle) = 0$ and $v(|\mathbf{y}\rangle) = 1$.

Stated differently [286, Theorem 2, p 183], let $|\mathbf{x}\rangle$ and $|\mathbf{y}\rangle$ be two non-orthogonal rays in a Hilbert space \mathfrak{H} of finite dimension ≥ 3 . Then there is a finite set of rays $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ containing $|\mathbf{x}\rangle$ and $|\mathbf{y}\rangle$ such that a (not necessarily two-valued) state v on $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ satisfies $v(|\mathbf{x}\rangle), v(|\mathbf{y}\rangle) \in \{0, 1\}$ only if $v(|\mathbf{x}\rangle) = v(|\mathbf{y}\rangle) = 0$. That is, if a system of three mutually exclusive outcomes (such as the spin of a spin-1 particle in a particular direction) is prepared in a definite state $|\mathbf{x}\rangle$ corresponding to $v(|\mathbf{x}\rangle) = 1$, then the state $v(|\mathbf{y}\rangle)$ along some direction $|\mathbf{y}\rangle$ which is neither collinear nor orthogonal to $|\mathbf{x}\rangle$ cannot be (pre-)determined, because, by an argument *via* some set of intertwined rays $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$, both cases would lead to a complete contradiction.

The proofs of the logical indeterminacy principle presented by Pitowsky and Hrushovski [286, 401] is global in the sense that any ray in the set of intertwining rays $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ in-between $|\mathbf{x}\rangle$ and $|\mathbf{y}\rangle$ – and thus not necessarily the “beginning and end points” $|\mathbf{x}\rangle$ and $|\mathbf{y}\rangle$ – may not have a pre-existing value. (If you are an omnirealist, substitute “pre-existing” by “non-contextual:” that is, any ray in the set of intertwining rays $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ may violate the admissibility rules and, in particular, non-contextuality.) Therefore, one might argue that the cases (i) as well as (ii); that is, $v(|\mathbf{x}\rangle) = v(|\mathbf{y}\rangle) = 1$, as well as $v(|\mathbf{x}\rangle) = 1$ and $v(|\mathbf{y}\rangle) = 0$ might still be predefined, whereas at least one ray in $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ cannot be pre-defined. (If you are an omnirealist, substitute “pre-defined” by “non-contextual.”)

This possibility has been excluded in a series of papers [3–6] *localizing value indefiniteness*. Thereby the strong admissibility rules coinciding with two-valued states which are total function on a logic, have been generalized or extended (if you prefer “weakened”) in such away as to allow for value definiteness. Essentially, by allowing the two-valued state to be a *partial* function on the logic, which need not be defined any longer on all of its elements, admissibility has been defined by two rules (IV) of Sect. 12.9.4: if $v(|\mathbf{x}\rangle) = 1$, then a measurement of all the other observables in a context containing $|\mathbf{x}\rangle$ must yield the value 0 for the other observables in this context – as well as counterfactually, in all contexts including $|\mathbf{x}\rangle$ and in mutually orthogonal rays which are orthogonal to $|\mathbf{x}\rangle$, such as depicted as the star-shaped configuration in Fig. 12.15. Likewise, if all propositions but one, say the one associated with $|\mathbf{x}\rangle$, in a context have value 0, then this proposition $|\mathbf{x}\rangle$ is assigned the value 1; that is, $v(|\mathbf{x}\rangle) = 1$.

However, as long as the entire context contains more than two atoms, if $v(|\mathbf{x}\rangle) = 0$ for some proposition associated with $|\mathbf{x}\rangle$, any of the other observables in the context containing $|\mathbf{x}\rangle$ could still yield the value 1 or 0. Therefore, these other observables need not be value definite. In such a formalism, and relative to the assumptions – in particular, by the admissibility rules allowing for value indefiniteness – sets of intertwined rays $\Gamma(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$ can be constructed which render value indefiniteness of

property $|y\rangle\langle y|$ if the system is prepared in state $|x\rangle$ (and thus $v(|x\rangle) = 1$). More specifically, sets of intertwined rays $\Gamma(|x\rangle, |y\rangle)$ can be found which demonstrate that, in accord with the “weak” admissibility rules (IV) of Sect. 12.9.4, in Hilbert spaces of dimension greater than two, in accord with complementarity, any proposition which is complementary with respect to the state prepared must be value indefinite [3–6].

How Can You Measure a Contradiction?

Clifton replied with this (rhetorical) question after I had asked if he could imagine any possibility to somehow “operationalize” the Kochen–Specker theorem.

Indeed, the Kochen–Specker theorem – in particular, not only non-separability but the total absence of any two-valued state – has been resilient to attempts to somehow “measure” it: first, as alluded by Clifton, its proof is by contraction – any assumption or attempt to consistently (in accordance with admissibility) construct two-valued state on certain finite subsets of quantum logics provably fails.

Second, the very absence of any two-valued state on such logics reveals the futility of any attempt to somehow define classical probabilities; let alone the derivation of any Boole’s conditions of physical experience – both rely on, or are, the hull spanned by the vertices derivable from two-valued states (if the latter existed) and the respective correlations. So, in essence, on logics corresponding to Kochen–Specker configurations, such as the I_2 -configuration of Kochen–Specker [314, p. 69], or the Cabello, Estebarez and García-Alcaine logic [91, 96] depicted in Fig. 12.16 which (subject to admissibility) have no two-valued states, classical probability theory breaks down entirely – that is, in the most fundamental way; by not allowing any two-valued state.

It is amazing how many papers exist which claim to “experimentally verify” the Kochen–Specker theorem. However, without exception, those experiments either prove some kind of Bell–Boole of inequality on single-particles (to be fair this is referred to as “proving contextuality;” such as, for instance, Refs. [36, 98, 267, 268, 309]); or show that the quantum predictions yield complete contradictions if one “forces” or assumes the counterfactual co-existence of observables in different contexts (and measured in separate, distinct experiments carried out in different subensembles; e.g., Refs. [91, 250, 383, 467, 468]; again these lists of references are incomplete.)

Of course, what one could still do is measuring all contexts, or subsets of compatible observables (possibly by Einstein–Podolsky–Rosen type [196] counterfactual inference) – one at a time – on different subensembles prepared in the same state by Einstein–Podolsky–Rosen type [196] experiments, and comparing the complete sets of results with classical predictions [250]. For instance, multiplying all products of dichotomic ± 1 observables within contexts, and summing up the results in parity proofs such as for the Cabello, Estebarez and García-Alcaine logic depicted in Fig. 12.16 must yield differences between the classical and the quantum predictions – in this case parity odd and even, respectively.

Contextual Inequalities

If one is willing to drop admissibility altogether while at the same time maintaining non-contextuality – thereby only assuming that the hidden variable theories

assign values to all the observables [54, Sect. 4, p. 375], thereby only assuming non-contextuality [92], one arrives at *contextual inequalities* [16]. Of course, these value assignments need to be much more general as the admissibility requirements on two-valued states; allowing all 2^n (instead of just n combinations) of contexts with n atoms; such as $1 - 1 - 1 - \dots - 1$, or $0 - 0 - \dots - 0$. For example, Cabello has suggested [92] to consider fourth order correlations within all the contexts (blocks; really within single maximal observables) constituting the logic considered by Cabello, Estebaranz and García-Alcaine [91, 96], and depicted as a Greechie orthogonality diagram in Fig. 12.16. For the sake of demonstration, consider a Greechie (orthogonality) diagram of a finite subset of the continuum of blocks or contexts imbeddable in four-dimensional real Hilbert space without a two-valued probability measure. More explicitly, the correlations are with nine tightly interconnected contexts $a = \{a_1, a_2, a_3, a_4\}$, $b = \{a_4, a_5, a_6, a_7\}$, $c = \{a_7, a_8, a_9, a_{10}\}$, $d = \{a_{10}, a_{11}, a_{12}, a_{13}\}$, $e = \{a_{13}, a_{14}, a_{15}, a_{16}\}$, $f = \{a_{16}, a_{17}, a_{18}, a_1\}$, $g = \{a_6, a_8, a_{15}, a_{17}\}$, $h = \{a_3, a_5, a_{12}, a_{14}\}$, $i = \{a_2, a_9, a_{11}, a_{18}\}$, respectively.

A hull problem can be defined as follows: (i) assume that each one of the 18 (partially counterfactual) observables a_1, a_2, \dots, a_{18} independently acquires either the definite value “ -1 ” or “ $+1$,” respectively. There are $2^{18} = 262144$ such cases. Note that, essentially, thereby all information on the intertwine structure is eliminated (the only remains are in the correlations taken in the next step), as one treats all observables to belong to a large Boolean algebra of 18 atoms a_1, a_2, \dots, a_{18} ; (ii) form all the 9 four-order correlations according to the context (block) structure $a_1a_2a_3a_4, a_4a_5a_6a_7, \dots, a_2a_9a_{11}a_{18}$, respectively; (iii) then evaluate (by multiplication) each one of these nine observables according to the valuations created in (i); (iv) for each one of the 2^{18} valuations form a 9-dimensional vector $(E_1 = a_1a_2a_3a_4, E_2 = a_4a_5a_6a_7, \dots, E_9 = a_2a_9a_{11}a_{18})^\top$ which contains all the values computed in (iii), and consider them as vertices (of course, there will be many duplicates which can be eliminated) defining a correlation polytope; (v) finally, solve the hull problem for this polytope. The resulting 274 inequalities and 256 vertices (a reverse vertex computation reveals 256 vertices; down from 2^{18}) confirms Cabello’s [92] as well as other bounds [521, Eq. (8)]; among them

$$\begin{aligned}
 & -1 \leq E_1 \leq 1, \\
 & E_1 + 7 \geq E_2 + E_3 + E_4 + E_5 + E_6 + E_7 + E_8 + E_9, \\
 & E_1 + E_8 + E_9 + 7 \geq E_2 + E_3 + E_4 + E_5 + E_6 + E_7, \\
 & E_1 + E_6 + E_7 + E_8 + E_9 + 7 \geq E_2 + E_3 + E_4 + E_5, \\
 & E_1 + E_4 + E_5 + E_6 + E_7 + E_8 + E_9 + 7 \geq E_2 + E_3, \\
 & E_1 + E_2 + E_3 + E_4 + E_5 + E_6 + E_7 + E_8 + E_9 + 7 \geq 0.
 \end{aligned} \tag{12.52}$$

Similar calculations for the pentagon and the Specker bug logics, by “bundling” the 3rd order correlations within the contexts (blocks, 3-atomic Boolean subalgebras), yield 32 (down from $2^{10} = 1024$ partially duplicate) vertices and 10 “trivial”

inequalities for the bug logic, as well as 128 (down from $2^{13} = 8192$ partially duplicate) vertices and 14 “trivial” inequalities for the Specker bug logic.

12.9.9 Quantum Probabilities and Expectations

Since from Hilbert space dimension higher than two there do not exist any two-valued states, the (quasi-)classical Boolean strategy to find (or define) probabilities *via* the convex sum of two-valued states brakes down entirely. Therefore, as this happened to be [172, 173, 295, 551, 552, 554], the quantum probabilities have to be “derived” or postulated from entirely new concepts, based upon quantities – such as vectors or projection operators – in linear vector spaces equipped with a scalar product. One guiding principle should be that, among those observables which are simultaneously co-measurable (that is, whose projection operators commute), the classical probability theory should hold.

Historically, what is often referred to as *Born rule* for calculating probabilities, has been a statistical re-interpretation of Schrödinger’s wave function [68, Footnote 1, Anmerkung bei der Korrektur, p. 865], as outlined by Dirac [172, 173] (a digression: a small piece [176] on “*the futility of war*” by the late Dirac is highly recommended; I had the honour listening to the talk personally), Jordan [295], von Neumann [551, 552, 554], and Lüders [89, 346, 347].

Rather than stating it as axiom of quantum mechanics, Gleason [240] derived the Born rule from elementary assumptions; in particular from subclassicality: within contexts – that is, among mutually commuting and thus simultaneously co-measurable observables – the quantum probabilities should reduce to the classical, Kolmogorovian, form. In particular, the probabilities of propositions corresponding to observables which are (i) mutually exclusive (in geometric terms: correspond to orthogonal vectors/projectors) as well as (ii) simultaneously co-measurable observables are (i) non-negative, (ii) normalized, and (iii) finite additive as in Eqs. (12.24) and (12.25); that is, probabilities (of atoms within contexts or blocks) add up to one [259, Sect. 1].

As already mentioned earlier, Gleason’s paper made a high impact on those in the community capable of comprehending it [41, 151, 180, 301, 314, 401, 434, 591]. Nevertheless it might not be unreasonable to state that, while a proof of the Kochen–Specker theorem is straightforward, Gleason’s results are less attainable. However, in what follows we shall be less concerned with either necessity nor with mixed states, but shall rather concentrate on sufficiency and pure states. (This will also rid us of the limitations to Hilbert spaces of dimensions higher than two.)

Recall that pure states [172, 173] as well as elementary yes-no propositions [62, 552, 554] can both be represented by (normalized) vectors in some Hilbert space. If one prepares a pure state corresponding to a unit vector $|\mathbf{x}\rangle$ (associated with the one-dimensional projection operator $\mathbf{E}_{\mathbf{x}} = |\mathbf{x}\rangle\langle\mathbf{x}|$) and measures an elementary yes-no proposition, representable by a one-dimensional projection operator $\mathbf{E}_{\mathbf{y}} = |\mathbf{y}\rangle\langle\mathbf{y}|$ (associated with the vector $|\mathbf{y}\rangle$), then Gleason notes [240, p. 885] in the second

paragraph that (in Dirac notation), “it is easy to see that such a [[probability]] measure μ can be obtained by selecting a vector $|\mathbf{y}\rangle$ and, for each closed subspace A , taking $\mu(A)$ as the square of the norm of the projection of $|\mathbf{y}\rangle$ on A .”

Since in Euclidean space, the projection \mathbf{E}_y of $|\mathbf{y}\rangle$ on $\mathfrak{A} = \text{span}(|\mathbf{x}\rangle)$ is the dot product (both vectors $|\mathbf{x}\rangle, |\mathbf{y}\rangle$ are supposed to be normalized) $|\mathbf{x}\rangle\langle\mathbf{x}|\mathbf{y}\rangle = |\mathbf{x}\rangle \cos\angle(|\mathbf{x}\rangle, |\mathbf{y}\rangle)$, Gleason’s observation amounts to the well-known quantum mechanical cosine square probability law referring to the probability to find a system prepared a in state in another, observed, state. (Once this is settled, all self-adjoint observables follow by linearity and the spectral theorem.)

In this line of thought, “measurement” contexts (orthonormal bases) allow “views” on “prepared” contexts (orthonormal bases) by the respective projections.

For the sake of demonstration, suppose some unit vector $|\rho\rangle$ corresponding to a pure quantum state (preparation) is selected. For each one-dimensional closed subspace corresponding to a one-dimensional orthogonal projection observable (interpretable as an elementary yes-no proposition) $E = |\mathbf{e}\rangle\langle\mathbf{e}|$ along the unit vector $|\mathbf{e}\rangle$, define $w_\rho(|\mathbf{e}\rangle) = |\langle\mathbf{e}|\rho\rangle|^2$ to be the square of the length $|\langle\rho|\mathbf{e}\rangle|$ of the projection of $|\rho\rangle$ onto the subspace spanned by $|\mathbf{e}\rangle$.

The reason for this is that an orthonormal basis $\{|\mathbf{e}_i\rangle\}$ “induces” an ad hoc probability measure w_ρ on any such context (and thus basis). To see this, consider the length of the orthogonal (with respect to the basis vectors) projections of $|\rho\rangle$ onto all the basis vectors $|\mathbf{e}_i\rangle$, that is, the norm of the resulting vector projections of $|\rho\rangle$ onto the basis vectors, respectively. This amounts to computing the absolute value of the Euclidean scalar products $\langle\mathbf{e}_i|\rho\rangle$ of the state vector with all the basis vectors.

In order that all such absolute values of the scalar products (or the associated norms) sum up to one and yield a probability measure as required in Eqs. (12.24) and (12.25), recall that $|\rho\rangle$ is a unit vector and note that, by the Pythagorean theorem, these absolute values of the individual scalar products – or the associated norms of the vector projections of $|\rho\rangle$ onto the basis vectors – must be squared. Thus the value $w_\rho(|\mathbf{e}_i\rangle)$ must be the square of the scalar product of $|\rho\rangle$ with $|\mathbf{e}_i\rangle$, corresponding to the square of the length (or norm) of the respective projection vector of $|\rho\rangle$ onto $|\mathbf{e}_i\rangle$. For complex vector spaces one has to take the absolute square of the scalar product; that is, $f_\rho(|\mathbf{e}_i\rangle) = |\langle\mathbf{e}_i|\rho\rangle|^2$.

Pointedly stated, from this point of view the probabilities $w_\rho(|\mathbf{e}_i\rangle)$ are just the (absolute) squares of the coordinates of a unit vector $|\rho\rangle$ with respect to some orthonormal basis $\{|\mathbf{e}_i\rangle\}$, representable by the square $|\langle\mathbf{e}_i|\rho\rangle|^2$ of the length of the vector projections of $|\rho\rangle$ onto the basis vectors $|\mathbf{e}_i\rangle$ – one might also say that each orthonormal basis allows “a view” on the pure state $|\rho\rangle$. In two dimensions this is illustrated for two bases in Fig. 12.18. The squares come in because the absolute values of the individual components do not add up to one; but their squares do. These considerations apply to Hilbert spaces of any, including two, finite dimensions. In this non-general, ad hoc sense the Born rule for a system in a pure state and an elementary proposition observable (quantum encodable by a one-dimensional projection operator) can be motivated by the requirement of additivity for arbitrary finite dimensional Hilbert space.

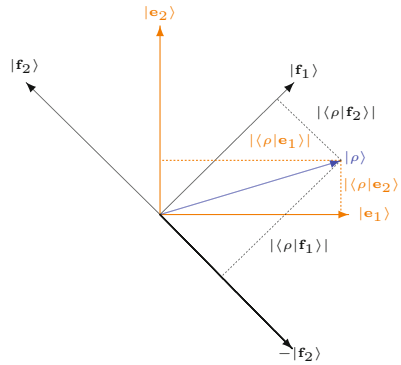


Fig. 12.18 Different orthonormal bases $\{|e_1\rangle, |e_2\rangle\}$ and $\{|f_1\rangle, |f_2\rangle\}$ offer different “views” on the pure state $|\rho\rangle$. As $|\rho\rangle$ is a unit vector it follows from the Pythagorean theorem that $|\langle\rho|e_1\rangle|^2 + |\langle\rho|e_2\rangle|^2 = |\langle\rho|f_1\rangle|^2 + |\langle\rho|f_2\rangle|^2 = 1$, thereby motivating the use of the absolute value (modulus) squared of the amplitude for quantum probabilities on pure states

12.9.9.1 Comparison of Classical and Quantum Form of Correlations

In what follows quantum configurations corresponding to the logics presented in the earlier sections will be considered. All of them have quantum realizations in terms of vectors spanning one-dimensional subspaces corresponding to the respective one-dimensional projection operators.

The appendix contains a detailed derivation of two-particle correlation functions. It turns out that, whereas on the singlet state the classical correlation function (B.1) $E_{c,2,2}(\theta) = \frac{2}{\pi}\theta - 1$ is linear, the quantum correlations (B.11) and (B.23) are of the “stronger” cosine form $E_{q,2j+1,2}(\theta) \propto -\cos(\theta)$. A stronger-than-quantum correlation would be a sign function $E_{s,2,2}(\theta) = \text{sgn}(\theta - \pi/2)$ [321].

When translated into the most fundamental empirical level – to two clicks in $2 \times 2 = 4$ respective detectors, a single click on each side – the resulting differences

$$\begin{aligned} \Delta E &= E_{c,2,2}(\theta) - E_{q,2j+1,2}(\theta) \\ &= -1 + \frac{2}{\pi}\theta + \cos\theta = \frac{2}{\pi}\theta + \sum_{k=1}^{\infty} \frac{(-1)^k \theta^{2k}}{(2k)!} \end{aligned} \quad (12.53)$$

signify a critical difference with regards to the occurrence of joint events: both classical and quantum systems perform the same at the three points $\theta \in \{0, \frac{\pi}{2}, \pi\}$. In the region $0 < \theta < \frac{\pi}{2}$, ΔE is strictly positive, indicating that quantum mechanical systems “outperform” classical ones with regard to the production of *unequal pairs* “+–” and “–+,” as compared to equal pairs “++” and “--.” This gets largest at $\theta_{\max} = \arcsin(2/\pi) \approx 0.69$; at which point the differences amount to 38% of all such pairs, as compared to the classical correlations. Conversely, in the region $\frac{\pi}{2} < \theta < \pi$, ΔE is strictly negative, indicating that quantum mechanical

systems “outperform” classical ones with regard to the production of *equal pairs* “++” and “--,” as compared to unequal pairs “+-” and “-+.” This gets largest at $\theta_{\min} = \pi - \arcsin(2/\pi) \approx 2.45$. Stronger-than-quantum correlations [414, 415] could be of a sign functional form $E_{s,2,2}(\theta) = \text{sgn}(\theta - \pi/2)$ [321].

In correlation experiments these differences are the reason for violations of Boole’s (classical) conditions of possible experience. Therefore, it appears not entirely unreasonable to speculate that the non-classical behaviour already is expressed and reflected at the level of these two-particle correlations, and not in need of any violations of the resulting inequalities.

12.9.10 Min-Max Principle

Violation of Boole’s (classical) conditions of possible experience by the quantum probabilities, correlations and expectations are indications of some sort of non-classicality; and are often interpreted as certification of quantum physics, and quantum physical features [395, 540]. Therefore it is important to know the extent of such violations; as well as the experimental configurations (if they exist [478]) for which such violations reach a maximum.

The basis of the min-max method are two observations [212]:

1. Boole’s bounds are *linear* – indeed linearity is, according to Pitowsky [400], the main finding of Boole with regards to *conditions of possible (nowadays classical physical) experience* [66, 67] – in the terms entering those bounds, such as probabilities and n th order correlations or expectations.
2. All such terms, in particular, probabilities and n th order correlations or expectations, have a quantum realization as self-adjoint transformations. As coherent superpositions (linear sums and differences) of self-adjoint transformations are again self-adjoint transformations (and thus normal operators), they are subject to the spectral theorem. So, effectively, all those terms are “bundled together” to give a single “comprehensive” (with respect to Boole’s conditions of possible experience) observable.
3. The spectral theorem, when applied to self-adjoint transformations obtained from substituting the quantum terms for the classical terms, yields an eigensystem consisting of all (pure or non-pure) states, as well as the associated eigenvalues which, according to the quantum mechanical axioms, serve as the measurement outcomes corresponding to the combined, bundled, “comprehensive,” observables. (In the usual Einstein–Podolsky–Rosen “explosion type” setup these quantities will be highly non-local.) The important observation is that this “comprehensive” (with respect to Boole’s conditions of possible experience) observable encodes or includes all possible one-by-one measurements on each one of the single terms alone, at least insofar as they pertain to Boole’s conditions.

4. By taking the minimal and the maximal eigenvalue in the spectral sum of this comprehensive observable one therefore obtains the minimal and the maximal measurement outcomes “reachable” by quantization.

Thereby, Boole’s conditions of possible experience are taken as given and for granted; and the computational intractability of their hull problem [399] is of no immediate concern, because nothing need to be said of actually finding those conditions of possible experience, whose calculation may grow exponential with the number of vertices. Note also that there might be a possible confusion of the term “min-max principle” [260, Sect. 90] with the term “maximal operator” [260, Sect. 84]. And finally, this is no attempt to compute general quantum ranges, as for instance discussed by Pitowsky [396, 402, 406] and Tsirelson [141–143].

Indeed, functional analysis provides a technique to compute (maximal) violations of Boole–Bell type inequalities [213, 214]: the *min-max principle*, also known as *Courant–Fischer–Weyl min-max principle* for self-adjoint transformations (cf. Ref. [260, Sect. 90], Ref. [430, pp. 75ff], and Ref. [528, Sect. 4.4, pp. 142ff]), or rather an elementary consequence thereof: by the spectral theorem any bounded self-adjoint linear operator \mathbf{T} has a spectral decomposition $\mathbf{T} = \sum_{i=1}^n \lambda_i \mathbf{E}_i$, in terms of the sum of products of bounded eigenvalues times the associated orthogonal projection operators. Suppose for the sake of demonstration that the spectrum is non-degenerate. Then we can (re)order the spectral sum so that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ (in case the eigenvalues are also negative, take their absolute value for the sort), and consider the greatest eigenvalue.

In quantum mechanics the maximal eigenvalue of a self-adjoint linear operator can be identified with the maximal value of an observation. Thereby, the spectral theorem supplies even the state associated with this maximal eigenvalue λ_1 : it is the eigenvector (linear subspace) $|\mathbf{e}_1\rangle$ associated with the orthogonal projector $\mathbf{E}_1 = |\mathbf{e}_1\rangle\langle\mathbf{e}_1|$ occurring in the (re)ordered spectral sum of \mathbf{T} .

With this in mind, computation of maximal violations of all the Boole–Bell type inequalities associated with Boole’s (classical) conditions of possible experience is straightforward:

1. take all terms containing probabilities, correlations or expectations and the constant real-valued coefficients which are their multiplicative factors; thereby excluding single constant numerical values $O(1)$ (which could be written on “the other” side of the inequality; resulting if what might look like “ $T(p_1, \dots, p_n, p_{1,2}, \dots, p_{123}, \dots) \leq O(1)$ ” (usually, these inequalities, for reasons of operationalizability, as discussed earlier, do not include higher than 2nd order correlations), and thereby define a function T ;
2. in the transition “quantization” step $T \rightarrow \mathbf{T}$ substitute all classical probabilities and correlations or expectations with the respective quantum self-adjoint operators, such as for two spin- $\frac{1}{2}$ particles enumerated in Eq. (B.6), $p_1 \rightarrow q_1 = \frac{1}{2} [\mathbb{I}_2 \pm \sigma(\theta_1, \varphi_1)] \otimes \mathbb{I}_2$, $p_2 \rightarrow q_2 = \frac{1}{2} [\mathbb{I}_2 \pm \sigma(\theta_2, \varphi_2)] \otimes \mathbb{I}_2$, $p_{12} \rightarrow q_{12} = \frac{1}{2} [\mathbb{I}_2 \pm \sigma(\theta_1, \varphi_1)] \otimes \frac{1}{2} [\mathbb{I}_2 \pm \sigma(\theta_2, \varphi_2)]$, $E_c \rightarrow \mathbf{E}_q = p_{12++} + p_{12--} - p_{12+-} - p_{12-+}$, as demanded by the inequality. Note that, since the coefficients in \mathbf{T} are all real-valued, and because $(A + B)^\dagger = A^\dagger + B^\dagger = (A + B)$

for arbitrary self-adjoint transformations A, B , the real-valued weighted sum \mathbf{T} of self-adjoint transformations is again self-adjoint.

3. Finally, compute the eigensystem of \mathbf{T} ; in particular the largest eigenvalue λ_{\max} and the associated projector which, in the non-degenerate case, is the dyadic product of the “maximal state” $|\mathbf{e}_{\max}\rangle$, or $\mathbf{E}_{\max} = |\mathbf{e}_{\max}\rangle\langle\mathbf{e}_{\max}|$.
4. In a last step, maximize λ_{\max} (and find the associated eigenvector $|\mathbf{e}_{\max}\rangle$) with respect to variations of the parameters incurred in step (ii).

The min-max method yields a feasible, constructive method to explore the quantum bounds on Boole’s (classical) conditions of possible experience. Its application to other situations is feasible. A generalization to higher-dimensional cases appears tedious but with the help of automated formula manipulation straightforward.

12.9.10.1 Expectations from Quantum Bounds

The quantum expectation can be directly computed from spin state operators. For spin- $\frac{1}{2}$ particles, the relevant operator, normalized to eigenvalues ± 1 , is

$$\mathbf{T}(\theta_1, \varphi_1; \theta_2, \varphi_2) = \left[2\mathbf{S}_{\frac{1}{2}}(\theta_1, \varphi_1) \right] \otimes \left[2\mathbf{S}_{\frac{1}{2}}(\theta_2, \varphi_2) \right]. \quad (12.54)$$

The eigenvalues are $-1, -1, 1, 1$ and 0 ; with eigenvectors for $\varphi_1 = \varphi_2 = \frac{\pi}{2}$,

$$\begin{aligned} &(-e^{-i(\theta_1+\theta_2)}, 0, 0, 1)^\top, \\ &(0, -e^{-i(\theta_1-\theta_2)}, 1, 0)^\top, \\ &(e^{-i(\theta_1+\theta_2)}, 0, 0, 1)^\top, \\ &(0, e^{-i(\theta_1-\theta_2)}, 1, 0)^\top, \end{aligned} \quad (12.55)$$

respectively.

If the states are restricted to Bell basis states $|\Psi^\mp\rangle = \frac{1}{\sqrt{2}}(|01\rangle \mp |10\rangle)$ and $|\Phi^\mp\rangle = \frac{1}{\sqrt{2}}(|00\rangle \mp |11\rangle)$ and the respective projection operators are \mathbf{E}_{Ψ^\mp} and \mathbf{E}_{Φ^\mp} , then the correlations, reduced to the projected operators $\mathbf{E}_{\Psi^\mp}\mathbf{E}\mathbf{E}_{\Psi^\mp}$ and $\mathbf{E}_{\Phi^\mp}\mathbf{E}\mathbf{E}_{\Phi^\mp}$ on those states, yield extrema at $-\cos(\theta_1 - \theta_2)$ for \mathbf{E}_{Ψ^-} , $\cos(\theta_1 - \theta_2)$ for \mathbf{E}_{Ψ^+} , $-\cos(\theta_1 + \theta_2)$ for \mathbf{E}_{Φ^-} , and $\cos(\theta_1 + \theta_2)$ for \mathbf{E}_{Φ^+} .

12.9.10.2 Quantum Bounds on the Clauser–Horne–Shimony–Holt Inequalities

The ease of this method can be demonstrated by (re)deriving the Tsirelson bound [141] of $2\sqrt{2}$ for the quantum expectations of the Clauser–Horne–Shimony–Holt inequalities (12.32) (cf. Sect. 12.9.8.2), which compare to the classical bound 2. First note that the two-particle projection operators along directions $\varphi_1 = \varphi_2 = \frac{\pi}{2}$ and θ_1, θ_2 ,

as taken from Eqs. (B.6) and (B.3), are

$$\begin{aligned} q_{1,\pm_1,2,\pm_2} \left(\theta_1, \varphi_1 = \frac{\pi}{2}, \theta_2, \varphi_2 = \frac{\pi}{2} \right) &= \\ &= \frac{1}{2} \left[\mathbb{I}_2 \pm \sigma \left(\theta_1, \frac{\pi}{2} \right) \right] \otimes \frac{1}{2} \left[\mathbb{I}_2 \pm \sigma \left(\theta_2, \frac{\pi}{2} \right) \right]. \end{aligned} \quad (12.56)$$

Adding these four orthogonal projection operators according to the parity of their signatures $\pm_1 \pm_2$ yields the expectation value

$$\begin{aligned} \mathbf{E}_q \left(\theta_1, \varphi_1 = \frac{\pi}{2}; \theta_2, \varphi_2 = \frac{\pi}{2} \right) &= \\ = \mathbf{E}_q(\theta_1, \theta_2) &= p_{1+2+} + p_{1-2-} - p_{1+2-} - p_{1-2+} = \\ &= \begin{pmatrix} 0 & 0 & 0 & e^{-i(\theta_1+\theta_2)} \\ 0 & 0 & e^{-i(\theta_1-\theta_2)} & 0 \\ 0 & e^{i(\theta_1-\theta_2)} & 0 & 0 \\ e^{i(\theta_1+\theta_2)} & 0 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (12.57)$$

Forming the Clauser–Horne–Shimony–Holt operator

$$\begin{aligned} \mathbf{CHSH}(\theta_1, \theta_2, \theta_3, \theta_4) &= \\ = \mathbf{E}_q(\theta_1, \theta_3) + \mathbf{E}_q(\theta_1, \theta_4) + \mathbf{E}_q(\theta_2, \theta_3) - \mathbf{E}_q(\theta_2, \theta_4). \end{aligned} \quad (12.58)$$

The eigenvalues

$$\begin{aligned} \lambda_{1,2} &= \mp 2\sqrt{1 - \sin(\theta_1 - \theta_2) \sin(\theta_3 - \theta_4)}, \\ \lambda_{3,4} &= \mp 2\sqrt{1 + \sin(\theta_1 - \theta_2) \sin(\theta_3 - \theta_4)}, \end{aligned} \quad (12.59)$$

for $\theta_1 - \theta_2 = \theta_3 - \theta_4 = \pm \frac{\pi}{2}$, yield the Tsirelson bounds $\pm 2\sqrt{2}$. In particular, for $\theta_1 = 0, \theta_2 = \frac{\pi}{2}, \theta_3 = \frac{\pi}{4}, \theta_4 = \frac{3\pi}{4}$, Eq. (12.58) reduces to

$$\begin{pmatrix} 0 & 0 & 0 & -2i\sqrt{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2i\sqrt{2} & 0 & 0 & 0 \end{pmatrix}; \quad (12.60)$$

and the eigenvalues are $\lambda_1 = 0, \lambda_2 = 0, \lambda_3 = -2\sqrt{2}, \lambda_4 = 2\sqrt{2}$; with the associated eigenstates $(0, 0, 1, 0)^\top, (0, 1, 0, 0)^\top, (i, 0, 0, 1)^\top, (-i, 0, 0, 1)^\top$, respectively. Note that, by comparing the components [368, p. 18] the eigenvectors associated with the eigenvalues reaching Tsirelson's bound are entangled, as could have been expected.

If one is interested in the measurements “along” Bell states, then one has to consider the projected operators $\mathbf{E}_{\psi^\mp}(\mathbf{CHSH})\mathbf{E}_{\psi^\mp}$ and $\mathbf{E}_{\phi^\mp}(\mathbf{CHSH})\mathbf{E}_{\phi^\mp}$ on those states which yield extrema at

$$\begin{aligned}
\lambda_{\psi_{\mp}} &= \mp \left[\cos(\theta_1 - \theta_3) + \cos(\theta_2 - \theta_3) + \right. \\
&\quad \left. + \cos(\theta_1 - \theta_4) - \cos(\theta_2 - \theta_4) \right], \\
\lambda_{\phi_{\mp}} &= \cos(\theta_1 + \theta_3) + \cos(\theta_2 + \theta_3) + \\
&\quad + \cos(\theta_1 + \theta_4) - \cos(\theta_2 + \theta_4).
\end{aligned} \tag{12.61}$$

For $\theta_1 = 0, \theta_2 = \frac{\pi}{2}, \theta_3 = \frac{\pi}{4}, \theta_4 = -\frac{\pi}{4}$, $\cos(\theta_1 + \theta_3) = \cos(\theta_2 + \theta_3) = \cos(\theta_1 + \theta_4) = -\cos(\theta_2 + \theta_4) = \frac{1}{\sqrt{2}}$, and Eq. (12.61) yields the Tsirelson bound $\lambda_{\psi_{\mp}} = \mp 2\sqrt{2}$. Likewise, for $\theta_1 = 0, \theta_2 = \frac{\pi}{2}, \theta_3 = -\frac{\pi}{4}, \theta_4 = \frac{\pi}{4}$, $\cos(\theta_1 + \theta_3) = \cos(\theta_2 + \theta_3) = \cos(\theta_1 + \theta_4) = -\cos(\theta_2 + \theta_4) = \frac{1}{\sqrt{2}}$, and Eq. (12.61) yields the Tsirelson bound $\lambda_{\phi_{\mp}} = \mp 2\sqrt{2}$.

Again it should be stressed that these violations might be seen as a “build-up;” resulting from the multiple addition of correlations which they contain.

Note also that, only as single context can be measured on a single system, because other context contain incompatible, complementary observables. However, as each observable is supposed to have the same (counterfactual) measurement outcome in any context, different contexts can be measured on different subensembles prepared in the same state such that, with the assumptions made (in particular, existence and context independence), Boole’s conditions of possible experience should be valid for the averages over each subensemble – regardless of whether they are co-measurable or incompatible and complementary. (This is true for instance for models with partition logics, such as generalized urn or finite automaton models.)

12.9.10.3 Quantum Bounds on the Pentagon

In a similar way two-particle correlations of a spin-one system can be defined by the operator \mathbf{S}_1 introduced in Eq. (B.13)

$$\mathbf{A}(\theta_1, \varphi_1; \theta_2, \varphi_2) = \mathbf{S}_1(\theta_1, \varphi_1) \otimes \mathbf{S}_1(\theta_2, \varphi_2). \tag{12.62}$$

Plugging in these correlations into the Klyachko–Can–Biniciogolu–Shumovsky inequality [312] in Eq. (12.40) yields the Klyachko–Can–Biniciogolu–Shumovsky operator

$$\begin{aligned}
&\mathbf{KCBS}(\theta_1, \dots, \theta_5, \varphi_1, \dots, \varphi_5) = \\
&= \mathbf{A}(\theta_1, \varphi_1, \theta_3, \varphi_3) + \mathbf{A}(\theta_3, \varphi_3, \theta_5, \varphi_5) + \mathbf{A}(\theta_5, \varphi_5, \theta_7, \varphi_7) + \\
&\quad + \mathbf{A}(\theta_7, \varphi_7, \theta_9, \varphi_9) + \mathbf{A}(\theta_9, \varphi_9, \theta_1, \varphi_1).
\end{aligned} \tag{12.63}$$

Taking the special values of Tkadlec [532], as enumerated in Cartesian coordinates in Fig. 12.6, which, in spherical coordinates, are $a_1 = (1, \frac{\pi}{2}, 0)^T$, $a_2 = (1, \frac{\pi}{2}, \frac{\pi}{2})^T$, $a_3 = (1, 0, \frac{\pi}{2})^T$, $a_4 = (\sqrt{2}, \frac{\pi}{2}, -\frac{\pi}{4})^T$, $a_5 = (\sqrt{2}, \frac{\pi}{2}, \frac{\pi}{4})^T$, $a_6 = (\sqrt{6}, \tan^{-1}(\frac{1}{\sqrt{2}}), -\frac{\pi}{4})^T$, $a_7 = (\sqrt{3}, \tan^{-1}(\sqrt{2}), \frac{3\pi}{4})^T$, $a_8 = (\sqrt{6}, \tan^{-1}(\sqrt{5}),$

$\tan^{-1}(\frac{1}{2})^\top$, $a_9 = (\sqrt{2}, \frac{3\pi}{4}, \frac{\pi}{2})^\top$, $a_{10} = (\sqrt{2}, \frac{\pi}{4}, \frac{\pi}{2})^\top$, yields eigenvalues of **KCBS** in

$$\{-2.49546, 2.2288, -1.93988, 1.93988, -1.33721, 1.33721, -0.285881, 0.285881, 0.266666\} \quad (12.64)$$

all violating Eq. (12.40).

12.9.10.4 Quantum Bounds on the Cabello, Estebaranz and García-Alcaine logic

As a final exercise we shall compute the quantum bounds on the Cabello, Estebaranz and García-Alcaine logic [91, 96] which can be used in a parity proof of the Kochen–Specker theorem in 4 dimensions, as depicted in Fig. 12.16 (where also a representation of the atoms as vectors in \mathbb{R}^4 suggested by Cabello [92, Fig. 1] is enumerated), as well as the dichotomic observables [92, Eq. (2)] $\mathbf{A}_i = 2|\mathbf{a}_i\rangle\langle\mathbf{a}_i| - \mathbb{I}_4$ is used. The observables are then “bundled” into the respective contexts to which they belong; and the context summed according to the contextual inequalities from the Hull computation (12.52), and introduced by Cabello [92, Eq. (1)]. As a result (we use Cabello’s notation and not ours),

$$\begin{aligned} \mathbf{T} = & -\mathbf{A}_{12} \otimes \mathbf{A}_{16} \otimes \mathbf{A}_{17} \otimes \mathbf{A}_{18} \\ & -\mathbf{A}_{34} \otimes \mathbf{A}_{45} \otimes \mathbf{A}_{47} \otimes \mathbf{A}_{48} - \mathbf{A}_{17} \otimes \mathbf{A}_{37} \otimes \mathbf{A}_{47} \otimes \mathbf{A}_{67} \\ & -\mathbf{A}_{12} \otimes \mathbf{A}_{23} \otimes \mathbf{A}_{28} \otimes \mathbf{A}_{29} - \mathbf{A}_{45} \otimes \mathbf{A}_{56} \otimes \mathbf{A}_{58} \otimes \mathbf{A}_{59} \\ & -\mathbf{A}_{18} \otimes \mathbf{A}_{28} \otimes \mathbf{A}_{48} \otimes \mathbf{A}_{58} - \mathbf{A}_{23} \otimes \mathbf{A}_{34} \otimes \mathbf{A}_{37} \otimes \mathbf{A}_{39} \\ & -\mathbf{A}_{16} \otimes \mathbf{A}_{56} \otimes \mathbf{A}_{67} \otimes \mathbf{A}_{69} - \mathbf{A}_{29} \otimes \mathbf{A}_{39} \otimes \mathbf{A}_{59} \otimes \mathbf{A}_{69} \end{aligned} \quad (12.65)$$

The resulting $4^4 = 256$ eigenvalues of \mathbf{T} have numerical approximations as ordered numbers $-6.94177 \leq -6.67604 \leq \dots \leq 5.78503 \leq 6.023$, neither of which violates the contextual inequality (12.52) and Ref. [92, Eq. (1)].

12.9.11 What Can Be Learned from These Brain Teasers?

When reading the book of Nature, she obviously tries to tell us something very sublime yet simple; but what exactly is it? As mentioned earlier it seems that often discussants approach this particular book not with *evenly-suspended attention* [224, 225] but with strong – even ideologic [144] or evangelical [589] – (pre)dispositions. This might be one of the reasons why Specker called this area “haunted” [482]. With these *provisos* we shall enter the discussion.

Already in 1935 – possibly based to the Born rule for computing quantum probabilities which differ from classical probabilities on a global scale involving complementary observables, and yet coincide within contexts – Schrödinger pointed out (cf. also Pitowsky [400, footnote 2, p. 96]) that [539, p. 327] “*at no moment does there exist an ensemble of classical states of the model that squares with the totality of quantum mechanical statements of this moment.*”⁴ This seems to be the gist of what can be learned from the quantum probabilities: they cannot be accommodated entirely within a classical framework.

What can be positively said? Quantum mechanics grant operational access merely to a single context (block, maximal observable, orthonormal basis, Boolean subalgebra); and for all that operationally matters, all observables forming that context can be simultaneously value definite. (It could formally be argued that an entire star of contexts intertwined in a “true” proposition must be value definite, as depicted in Fig. 12.15.) A single context represents the maximal information encodable into a quantum system. This can be done by state preparation.

Beyond this single context one can get “views” on that single state in which the quantized system has been prepared. But these “views” come at a price: value indefiniteness. (Value indefiniteness is often expressed as “contextuality,” but this view is distractive, as it suggests some existing entity which is changing its value; depending on how – that is, along which context – it is measured.)

This situation might not be taken as a metaphysical conundrum, but perceived rather Socratically: it should come as no surprise that intrinsic [500], emdedded [538] observers have no access to all the information they subjectively desire, but only to a limited amount of properties their system – be it a virtual or a physical universe – is capable to express. Therefore there is no omniscience in the wider sense of “all that observers want to know” but rather than “all that is operationally realizable.”

Anything beyond this narrow “local omniscience covering a single context” in which the quantized system has been prepared appears to be a subjective illusion which is only stochastically supported by the quantum formalism – in terms of Gleason’s “projective views” on that single, value definite context. Experiments may enquire about such value indefinite observables by “forcing” a measurement upon a system not prepared or encoded to be “interrogated” in that way. However, these “measurements” of non-existing properties, although seemingly possessing viable outcomes which might be interpreted as referring to some alleged “hidden” properties, cannot carry any (consistent classical) content pertaining to that system alone.

To paraphrase a *dictum* by Peres [388], unprepared contexts do not exist; at least not in any operationally meaningful way. If one nevertheless forces metaphysical existence upon (value) indefinite, non-existing, physical entities the price, hedged into the quantum formalism, is stochasticity.

⁴German original [452, p. 811]: “*Es gibt in keinem Augenblick ein Kollektiv klassischer Modellzustände, auf das die Gesamtheit der quantenmechanischen Aussagen dieses Augenblicks zutrifft.*”

12.10 Quantum Mechanical Observer–Object Theory

The quantum measurement problem is at the heart of today’s quantum random number generators. Thus everybody relying on that technology has to be concerned with this seemingly philosophical issues of observer-object relation. And anybody denying its existence (aka Austin Powers’ “*if you got an issue here’s a tissue*”) is tantamount to building a bridge with a new material whose properties and construction objectives are largely unknown – thereby relying on assurances of most experts which are solely based on heuristics.

Presently the quantum mechanical observer—object theory is a “canvas with many facets and nuances.” There is no one accepted view of the measurement problem. The *Ansätze* proposed include, but are not limited to

- (I) collapse models: modification of quantum mechanics by the inclusion of some additional non-linear, irreversible transformation accounting for von Neumann’s process 1, and possible also for the transformation of pure states into mixed ones [226, 237, 238, 544].
- (II) Exner-Schrödinger thesis: all laws; in particular, also the unitary time evolution of the quantum state, have to be understood merely statistically, and are not valid individually [209, 262, 451].
- (III) Noncollapse Schrödinger-type quantum jellification without observation or measurement: in Schrödinger’s own words [457, pp. 19,20], “*He [[the quantum physicist]] thinks that if the laws of nature took this [[von Neumann’s Process 2, permutation]] form for, let me say, a quarter of an hour, we should find our surroundings rapidly turning into a quagmire, or sort of a featureless jelly or plasma, all contours becoming blurred, we ourselves probably becoming jelly fish. . . . nature is prevented from rapid jellification only by our perceiving or observing it. And I wonder that he is not afraid, when he puts a ten-pound-note {his wrist-watch} into his drawer in the evening, he might find it dissolved in the morning, because he has not kept watching it.*”
- (IV) Non-collapse Everettian type relative state formalism [30, 33, 205, 206, 208, 544, 545]: Everett realized that, due to nesting, there cannot occur any kind of irreversibility, but just the formation of entanglement. Entanglement in turn induces relational properties between objects and observers (a the price of individual properties of those parties). Given a pure state and a measurement not compatible with it, the coherent superposition decomposition of the state in terms of the eigenstates of the measurement operator all constitute a valid observer (observing agent) who subjectively experiences the outcome of the measurement as unique. Maybe Everett would have even granted observers the capacity of being in a coherent superposition yet having the illusory subjective experience of uniqueness; but this is highly speculative; as well as all further interpretations of his writings in terms of *splitting worlds theory* [207].
- (V) Non-collapse and intrinsic incompleteness: already von Neumann mentions (and immediately discards this as a solution of the measurement problem) the possibility that [554, Sect. 6.2, p. 426] “. . . *the result of the measurement is*

*indeterminate, because the state of the observer before the measurement is not known exactly. It is conceivable that such a mechanism might function, because the state of information of the observer regarding his own state could have absolute limitations, by the laws of nature.*⁵ Breuer has discussed this possibility in a series of papers [72–74]. This is not dissimilar to self-nesting, as discussed in Sect. 1.8.

(VI) Non-collapse entanglement (zero sum scenario, excluding consciousness): the extrinsic state representation of the combined object and observer system is pure and entangled, while intrinsically both the observer and the object, mistakenly perceived individually, are in mixed states. This line of thought might have been best expressed by London and Bauer [341, 342] who base their presentation on von Neumann’s treatment of the measurement process [552, Chap. VI] (echoed also in Everett’s [206] and Wigner’s [571] papers). Related ideas can be found in Schrödinger’s accounts on entanglement [452, 453, 455], which in turn have been influenced by a (nowadays famous) paper by Einstein, Podolsky and Rosen [196].

(i) Initial phase: this conceptualization of the measurement process starts with the supposition that initially the entire system consists of two isolated systems O and A , identified with an object and the measurement apparatus, respectively. Initially, if the respective states are pure and denoted by $|\psi_O\rangle$ as well as $|\psi_A\rangle$, then the wave function of the entire system can be composed from the individual parts by multiplication; that is, $|\psi_{O\&A}\rangle = |\psi_O\rangle \otimes |\psi_A\rangle = |\psi_O\psi_A\rangle$. In this initial phase the object as well as the measurement apparatus are separated and know nothing about each other, their joint state $|\psi_{O\&A}\rangle$ being non-entangled and without any relational properties.

Suppose further, for the sake of simplicity, that both the object as well as the measurement device have an equal number k of mutually exclusive states $|\psi_{O,i}\rangle$ as well as $|\psi_{A,j}\rangle$, with $1 \leq i, j \leq k$, respectively. The operator \mathbf{A} corresponding to the measurement device should have a spectral resolution of $\mathbf{E} = \sum_{j=1}^k a_j |\psi_{A,j}\rangle \langle \psi_{A,j}|$.

(ii) Interaction phase: in order to obtain information about each other, both object and the measurement apparatus have to interact with each other. This interaction is supposed to be representable by a unitary transformation. During this interaction, the initial state $|\psi_{O\&A}\rangle$ is transformed into a coherent superposition, a sum of products of individual states of O and A , so that the state after the interaction phase is $|\psi'_{O\&A}\rangle = \sum_{i,j=1}^k \varphi_{ij} |\psi_{O,i}\rangle \otimes |\psi_{A,j}\rangle = \sum_{i,j=1}^k \varphi_{ij} |\psi_{O,i}\psi_{A,j}\rangle$. Preferably, in a measurement, states of the measurement apparatus “should get aligned” with states of the object, such that $\varphi_{ij} \approx \delta_{ij} \varphi_{ii}$ and $|\psi'_{O\&A}\rangle \approx \sum_{i=1}^k \varphi_{ii} |\psi_{O,i}\psi_{A,i}\rangle$.

⁵German original [554, Sect. 6.3, p 233], “. . . das Resultat der Messung ist unbestimmt, weil der Zustand des Beobachters vor der Messung nicht genau bekannt ist. Es wäre denkbar, da ein solcher Mechanismus funktioniert, denn die Informiertheit des Beobachters über den eigenen Zustand könnte naturgesetzliche Schranken haben.”

This state $|\psi'_{O\&A}\rangle$ will in generally not be factorizable as a non-entangled product state of individual states of O and A . Suppose further that it is not the case; that is, suppose that $|\psi_{O\&A}\rangle$ is entangled.

- (iii) interpretation phase: depending on our viewpoint, conventions and inclinations,
- i. on the one hand, $|\psi_{O\&A}\rangle$ can either be perceived from the outside – that is, extrinsically – and thus appear as pure entangled state of the combined system of object and the measurement apparatus, encoding relational information (that is, statistical correlations) among these subsystems, but lacking complete information of individual subsystems;
 - ii. or, on the other hand, from the intrinsic point of view of individual subsystems, $|\psi_{O\&A}\rangle$ can be analysed in terms of the individual components by forcing some type of individuality (e.g., by taking the partial trace with respect to one subsystem) upon the subsystems. In the latter case of *individuality forcing*, as the wave function lacks complete information about the individual subsystems, the respective undefined subsystem properties are value indefinite.

As a side effect of individuality forcing, entanglement is fapp destroyed, and $|\psi'_{O\&A}\rangle$ undergoes a change back to a non-entangled state $|\psi''_{O\&A}\rangle$, subject to the relational information contained in $|\psi'_{O\&A}\rangle$.

One may ask how individuality can be enforced upon $|\psi'_{O\&A}\rangle$. This can be done by *context translation* [505] (for related ideas see Refs. [333, 520]); that is, by “translating” or “transforming” a misaligned measurement context into one which can be analysed, possibly by a third measurement device. This translation process introduces stochasticity through the (supposedly many) degrees of freedom of the outside measurement apparatus. Thereby, context translation involves fapp irreversibility [348] for macroscopic measurement devices [202, 461]. Yet in principle the chaining or nesting of measurements results in a (potentially infinite) regress.

If there is a regress, when does it stop? This can be answered by considering the smallest isolated system encompassing the original object O as well as the measurement apparatus A . Already Baumann [37] and Zeh [586] have pointed out that, strictly speaking, because of the high density of energy niveaus in macroscopic systems, the interactions between macroscopic systems are effective even at astronomical distances. Therefore these systems are exceedingly difficult to isolate; and any system which includes all (nested) observers would encompass the universe as a whole.

Most importantly, whatever the measurement outcome of measurements on individual parts of the entangled object–measurement apparatus system, this cannot correspond to some pre-existing, definite value solely residing within the bounds of the observed object, because the information encoded in the entangled state is (also, and in the extreme case solely and exclusively) in the

relational properties of its constituent parts; and not about the individual states of these constituents. Therefore, if one forces individuality, one has to add additional information from the environment, in particular, the measurement apparatus, which is not present in the original state.

The author is inclined to adopt this view of the measurement process – it is just like zooming in and out of the situation: if one looks at it from an extrinsic, outside, disentangled perspective (if one can afford such a view) – that is, as an isolated holistic system including the observer and the object, as well as the cut between them, the system is in a pure, well-defined state. However, if one “zooms into” this system, and takes an embedded, intrinsic point of view, then the individual constituents of the system – in particular, the object as well as the observer – are underdefined and value definite. “Forcing individuality” upon these constituents requires additional input from the environment (*via* context translation), thereby introducing auxiliary bits which do not reflect any property of those constituents.

- (VII) consciousness causes state reduction: this scenario is identical to the previous one but employs nesting until the level of consciousness of the observer. At this level, awareness by consciousness is then assumed to be essentially irreversible. That is, it is assumed that one cannot “unthink” the perception of a measurement. This point of view has been suggested by London and Bauer [341, 342] as well as Wigner [571]; although the latter one may have developed a different stance on this subject later [203]. For a critical discussion, see Ref. [582].

12.11 Observer-Objects “Riding” on the Same State Vector

What does it mean “to ride on a particular section” of a vector in high dimensional Hilbert space? Can two such sectors of one and the same vector constitute an observer-object system? Where is the cut, the interface between those sections?

We suggest here that indeed it might fapp be possible to make a distinction between observer and object; where both parties “ride” the same state. This distinction is formally specified by Everett’s relative states; that is, it involves entangled states.

12.12 Metaphysical Status of Quantum Value Indefiniteness

What does it mean that a particular (quantum) entity is value indefinite? It means that relative to, or with respect to, a particular physical resource or physical means, the respective entity cannot be entirely, that is, completely and totally, defined. In short: any proposition about physical value indefiniteness is means relative.

If such an entity is “observed” nevertheless, then this “observation” must necessarily introduce, add, input, and include, other specifications “outside” of the object.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

