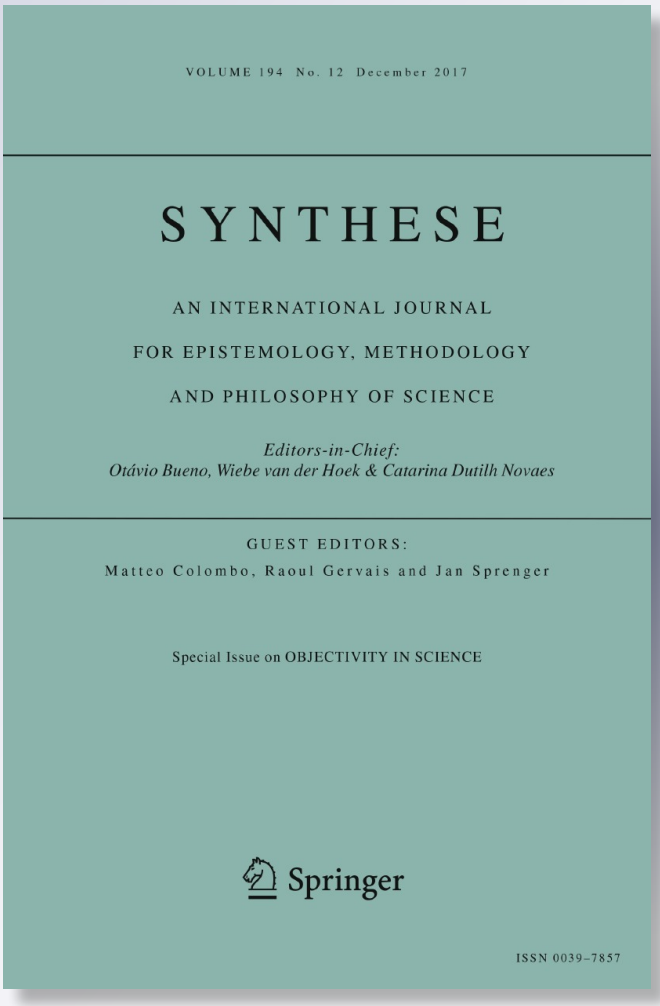
Quantum mechanics over sets: a pedagogical model with non-commutative finite probability theory as its quantum probability calculus

David Ellerman

Synthese
An International Journal for Epistemology, Methodology and Philosophy of Science

ISSN 0039-7857
Volume 194
Number 12

Synthese (2017) 194:4863-4896
DOI 10.1007/s11229-016-1175-0



Springer

Quantum Mechanics over Sets: A pedagogical model with non-commutative finite probability theory as its quantum probability calculus

David Ellerman
Department of Philosophy
U. of California/Riverside

August 8, 2016

Abstract

This paper shows how the classical finite probability theory (with equiprobable outcomes) can be reinterpreted and recast as the quantum probability calculus of a pedagogical or toy model of *quantum mechanics over sets* (QM/sets). There have been several previous attempts to develop a quantum-like model with the base field of $\mathbb{C}$ replaced by $\mathbb{Z}_2$. Since there are no inner products on vector spaces over finite fields, the problem is to define the Dirac brackets and the probability calculus. The previous attempts all required the brackets to take values in $\mathbb{Z}_2$. But the usual QM brackets $\langle\psi|\varphi\rangle$ give the "overlap" between states ψ and φ, so for subsets $S, T \subseteq U$, the natural definition is $\langle S|T\rangle = |S \cap T|$ (taking values in the natural numbers). This allows QM/sets to be developed with a full probability calculus that turns out to be a non-commutative extension of classical Laplace-Boole finite probability theory. The pedagogical model is illustrated by giving simple treatments of the indeterminacy principle, the double-slit experiment, Bell's Theorem, and identical particles in QM/Sets. A more technical appendix explains the mathematics behind carrying some vector space structures between QM over $\mathbb{C}$ and QM/Sets over $\mathbb{Z}_2$.

Contents

1 Introduction

This paper develops a pedagogical or "toy" model of quantum mechanics over sets (QM/sets) where the quantum probability calculus is a non-commutative version of the ordinary Laplace-Boole finite logical probability theory ([14], [3]) and where the usual vector spaces over $\mathbb{C}$ for QM are replaced with vector spaces over $\mathbb{Z}_2$ in QM/sets. Quantum mechanics over sets is a bare-bones "logical" (e.g., non-physical[1]) version of QM with appropriate versions of spectral decomposition, the Dirac brackets, the norm, observable-attributes, the Born rule, commutators, and density matrices all in the simple classical setting of sets, but that nevertheless provides models of characteristically quantum results (e.g., a QM/sets version of the indeterminacy principle, the double-slit experiment, Bell's Theorem, and the statistics for identical particles). In that manner, QM/sets can serve not only as a pedagogical (or "toy") model of QM but perhaps as an engine to better elucidate QM itself by representing the quantum features in a simple setting.

There have been at least three previous attempts at developing a version of QM where the base field of $\mathbb{C}$ is replaced by $\mathbb{Z}_2$ ([18], [12], and [21]). Since there are no inner products in vector spaces over a finite field, the "trick" is how to define the brackets, the norm, and then the probability algorithm. All these previous attempts use the aspect of full QM that the bras are dual vectors so the brackets take their values in the base field of $\mathbb{Z}_2$. For instance, the Schumacher-Westmoreland model does "not make use of the idea of probability"[18, p. 919] and have instead only a modal interpretation (1 = possibility and 0 = impossibility). There is a fourth category-theoretic model where the objects are sets [1] but it also has the "brackets" taking $0, 1$ values.[2]

The model of QM over sets developed here does not have the brackets taking values in the base field of $\mathbb{Z}_2$. It is based on a different understanding of the relation between the pedagogical or toy model and full QM. Instead of trying to mimic QM (replacing $\mathbb{C}$ with $\mathbb{Z}_2$), the idea is that QM/sets can perfectly well have the brackets and observables take values *outside* the base field of $\mathbb{Z}_2$ (e.g., use real-valued observables = real-valued random variables in classical finite probability theory) and even defining a more primitive version of "eigenvectors" and "eigenvalues" that are not (in general) the eigenvectors and eigenvalues of linear operators on the vector space over $\mathbb{Z}_2$. The transitioning from QM/sets to full QM is then seen *not* as going from one model to another model of a set of axioms (e.g., as in [1]) but as a process of "internalization" allowed by increasing the base field from $\mathbb{Z}_2$ to $\mathbb{C}$. The increased power of $\mathbb{C}$ (e.g., algebraic completeness) then allows the "eigenvectors" and "eigenvalues" of QM/sets (defined for arbitrary real-valued functions on a finite set) to be "internalized" as true eigenvectors and eigenvalues of (Hermitian) linear operators on vector spaces over $\mathbb{C}$ and the brackets can then also be "internalized" as a bilinear inner product taking values in the base field $\mathbb{C}$. Hence under this approach (and in contrast to the previous approaches), the

[1] In full QM, the DeBroglie relations connect mathematical notions such as frequency and wave-length to physical notions such as energy and momentum. QM/sets is "non-physical" in the sense that it is a sets-version of the pure mathematical framework of (finite-dimensional) QM without those direct physical connections.

[2] The Spekkens toy model [19] does not use vector spaces at all or utilize sets so it is not directly comparable.

"taking values in the base field" is seen *only* as an aspect of full QM over $\mathbb{C}$ and not as a necessary aspect of a pedagogical proto-QM model such as QM/sets with the base field of $\mathbb{Z}_2$.

What is the criterion of success for a toy model? Is the probability calculus another weird theory without interpretation that resembles quantum mechanics? Here the "proof of the pudding" is that the probability calculus of QM/sets *is* the classical Laplace-Boole finite probability theory–which is thereby extended to a non-commutative theory allowed by the vector space formulation.[3] Hence the pedagogical model allows a wide range of quantum phenomenon to be displayed in a rather simple setting.

2 Laplace-Boole finite probability theory

Since our purpose is conceptual rather than mathematical, we will stick to the simplest case of finite probability theory with a finite sample space or outcome space $U = \{u_1, ..., u_n\}$ of n equiprobable outcomes and to finite dimensional QM.[4] The *events* are the subsets $S \subseteq U$, and the *probability* of an event S occurring in a trial is the ratio of the cardinalities: $\Pr(S) = \frac{|S|}{|U|}$. Given that a conditioning event $S \subseteq U$ occurs, the *conditional probability* that $T \subseteq U$ occurs is: $\Pr(T|S) = \frac{\Pr(T \cap S)}{\Pr(S)} = \frac{|T \cap S|}{|S|}$. The ordinary probability $\Pr(T)$ of an event T can be taken as the conditional probability with U as the conditioning event so all probabilities can be seen as conditional probabilities. Given a (real-valued) random variable, here called an *attribute* $f : U \to \mathbb{R}$ on the elements of U, the *probability of observing a value r given an event S* is the conditional probability of the event $f^{-1}(r)$ given S:

$$\Pr(r|S) = \frac{|f^{-1}(r) \cap S|}{|S|}.$$

That is all the probability theory we will need here. Our task is to show how the mathematics of finite probability theory can be recast using the mathematical notions of quantum mechanics with the base field of $\mathbb{Z}_2$.

3 Recasting finite probability theory as a quantum probability calculus

3.1 Vector spaces over $\mathbb{Z}_2$

To show how classical Laplace-Boole finite probability theory can be recast as a quantum probability calculus, we use finite dimensional vector spaces over $\mathbb{Z}_2$. The power set $\wp(U)$ of $U = \{u_1, ..., u_n\}$ is a vector space over $\mathbb{Z}_2 = \{0, 1\}$, isomorphic to $\mathbb{Z}_2^n$, where the vector addition $S + T$ is the *symmetric difference* (exclusive-or on members) of subsets. That is, for $S, T \subseteq U$,

$$S + T = (S - T) \cup (T - S) = S \cup T - S \cap T$$

so the members of $S + T$ are the elements that are members of S or members of T but not members of both.

The *U-basis* in $\wp(U)$ is the set of singletons $\{u_1\}, \{u_2\}, ..., \{u_n\}$. A vector $S \in \wp(U)$ is specified in the U-basis as $S = \sum_{u \in S} \{u\}$ and it is characterized by its $\mathbb{Z}_2$-valued characteristic function $\chi_S : U \to \mathbb{Z}_2 \subseteq \mathbb{R}$ of coefficients since $S = \sum_{u \in U} \chi_S(u) \{u\}$. Similarly, a vector v in $\mathbb{C}^n$ is specified

[3] Instead of the fixed outcome set of classical probability theory, there is a vector space where each basis set plays the role of an outcome set or sample space. Since there are very different "incompatible" basis sets, "noncommutative" aspects of full QM appear in QM/sets.

[4] The mathematics can be generalized to the case where each point u_i in the sample space has a probability p_i but the simpler case of equiprobable points serves our conceptual purposes.

in terms of an orthonormal basis $\{|v_i\rangle\}$ as $v = \sum_i c_i |v_i\rangle$ and is characterized by a $\mathbb{C}$-valued function $\langle \_|v\rangle : \{v_i\} \to \mathbb{C}$ assigning a complex amplitude $\langle v_i|v\rangle = c_i$ to each basis vector $|v_i\rangle$.

Seeing $\wp(U)$ as the abstract vector space $\mathbb{Z}_2^n$ allows different bases in which the vectors can be expressed (as well as the basis-free notion of a vector as a "ket"). Hence the quantum probability calculus developed here can be seen as a "non-commutative" generalization of the classical Laplace-Boole finite probability theory where a different basis corresponds to a different equicardinal sample space $U' = \{u'_1, ..., u'_n\}$.

Consider the simple case of $U = \{a, b, c\}$ where the U-basis is $\{a\}$, $\{b\}$, and $\{c\}$. The three subsets $\{a, b\}$, $\{b, c\}$, and $\{a, b, c\}$ also form a basis since:

$\{b, c\} + \{a, b, c\} = \{a\}$;

$\{b, c\} + \{a, b\} + \{a, b, c\} = \{b\}$; and

$\{a, b\} + \{a, b, c\} = \{c\}$.

These new basis vectors could be considered as the basis-singletons in another equicardinal universe $U' = \{a', b', c'\}$ where $\{a'\}$, $\{b'\}$, and $\{c'\}$ refer to the same abstract vector as $\{a, b\}$, $\{b, c\}$, and $\{a, b, c\}$ respectively.

In the following *ket table*, each row is an abstract vector of $\mathbb{Z}_2^3$ expressed in the U-basis, the U'-basis, and a U''-basis.

$U = \{a, b, c\}$	$U' = \{a', b', c'\}$	$U'' = \{a'', b'', c''\}$
$\{a, b, c\}$	$\{c'\}$	$\{a'', b'', c''\}$
$\{a, b\}$	$\{a'\}$	$\{b''\}$
$\{b, c\}$	$\{b'\}$	$\{b'', c''\}$
$\{a, c\}$	$\{a', b'\}$	$\{c''\}$
$\{a\}$	$\{b', c'\}$	$\{a''\}$
$\{b\}$	$\{a', b', c'\}$	$\{a'', b''\}$
$\{c\}$	$\{a', c'\}$	$\{a'', c''\}$
$\emptyset$	$\emptyset$	$\emptyset$

Ket table giving a vector space isomorphism: $\mathbb{Z}_2^3 \cong \wp(U) \cong \wp(U') \cong \wp(U'')$ where row = ket.

In the Dirac notation [6], the *ket* $|\{a, c\}\rangle$ represents the abstract vector that is represented in the U-basis as $\{a, c\}$. A row of the ket table gives the different representations of the *same* ket in the different bases, e.g., $|\{a, c\}\rangle = |\{a', b'\}\rangle = |\{c''\}\rangle$.

3.2 The brackets and norm

In a Hilbert space, the inner product is used to define the brackets $\langle v_i|v\rangle$ and the norm $\|v\| = \sqrt{\langle v|v\rangle}$ but there are no inner products in vector spaces over finite fields. The different attempts to develop a toy model of QM over a finite field ([18], [21], [12]) such as $\mathbb{Z}_2$ differ from this model in how they address this problem. The treatment of the Dirac brackets and norm defined here is distinguished by the fact that the resulting probability calculus in QM/Sets is (a non-commutative version of) classical finite probability theory (instead of just a modal calculus with values 0 and 1).

For a singleton basis vector $\{u_j\} \subseteq U$, the (basis-dependent) *bra* $\langle \{u_j\}|_U : \wp(U) \to \mathbb{R}$ is defined by the *bracket*:

$$\langle \{u\}|_U S\rangle = \begin{cases} 1 \text{ if } u \in S \\ 0 \text{ if } u \notin S \end{cases} = |\{u_j\} \cap S| = \chi_S(u_j).$$

Note that the bra and the bracket is defined in terms of the U-basis and that is indicated by the U-subscript on the bra portion of the bracket. Then for $u_j, u_k \in U$, $\langle \{u_j\}|_U \{u_k\}\rangle = \chi_{\{u_k\}}(u_j) = \chi_{\{u_j\}}(u_k) = \delta_{jk}$ (the Kronecker delta function) which is the QM/Sets-version of $\langle v_j|v_k\rangle = \delta_{jk}$ for

an orthonormal basis $\{|v_j\rangle\}$ of $\mathbb{C}^n$. The bracket linearly extends *in the natural numbers* $\mathbb{N} \subseteq \mathbb{R}$ to any two vectors $T, S \in \wp(U)$:[5]

$$\langle T|_U S\rangle = |T \cap S|.$$

This is the QM/Sets-version of the Dirac brackets in the mathematics of QM.

This treatment of the brackets is motivated by the general method for transporting basis-set-defined structures between vector spaces over different fields, e.g., from $\mathbb{C}^n$ to $\mathbb{Z}_2^n$ (see Appendix). In both cases, the bracket gives a measure of the overlap or indistinctness of the two vectors.[6] The ket $|S\rangle$ is the same as the ket $|S'\rangle$ for some subset $S' \subseteq U'$ in another U'-basis, but when the bra $\langle\{u_j\}|_U$ is applied to the ket $|S\rangle = |S'\rangle$, then it is the subset $S \subseteq U$, not $S' \subseteq U'$, that comes outside the ket symbol $|\ \rangle$ in $\langle\{u_j\}|_U S\rangle = |\{u_j\} \cap S|$.[7] Heuristically, the bra $\langle T|_U$ can be thought of as a row-vector of zeros and ones expressed in the U-basis, and then the ket $|S\rangle$ is expressed as a column vector in the U-basis, and $\langle T|_U S\rangle$ is their dot product computed in the natural numbers with the usual embedding in the reals.

The *U-norm* $\|S\|_U : \wp(U) \to \mathbb{R}$ is defined, as usual, as the square root of the bracket:[8]

$$\|S\|_U = \sqrt{\langle S|_U S\rangle} = \sqrt{|S \cap S|} = \sqrt{|S|}$$

for $S \in \wp(U)$ which is the QM/Sets-version of the norm $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$ in ordinary QM. Hence $\|S\|_U^2 = |S|$ is the counting measure on $\wp(U)$. Note that a ket has to be expressed in the U-basis to apply the U-norm definition so, for example, $\|\{a'\}\|_U = \sqrt{2}$ since $|\{a'\}\rangle = |\{a, b\}\rangle$.

3.3 Numerical attributes and linear operators

In classical physics, the observables are numerical attributes, e.g., the assignment of a position and momentum to particles in phase space. One of the differences between classical and quantum physics is the replacement of these observable numerical attributes by linear operators associated with the observables where the values of the observables appear as eigenvalues of the operators. But this difference may be smaller than it would seem at first since a numerical attribute $f : U \to \mathbb{R}$ can be recast into an operator-like format in QM/sets, and there is even a QM/sets-analogue of spectral decomposition.

An observable, i.e., a Hermitian operator, on a finite-dimensional Hilbert space V has a home basis set of orthonormal eigenvectors. In a similar manner, a real-valued attribute $f : U \to \mathbb{R}$ defined on U has the U-basis as its "home basis set." The connection between the numerical attributes $f : U \to \mathbb{R}$ of QM/sets and the Hermitian operators of full QM can be established by seeing the function f as being *like* an "operator" $f \restriction ()$ on $\wp(U)$ in that it is used to define a sets-version of an "eigenvalue" equation [where $f \restriction S$ is the *restriction* of f to $S \in \wp(U)$]. For any subset $S \in \wp(U)$, the definition of the equation is:

[5] Here $\langle T|_U S\rangle = |T \cap S|$ takes values in the natural numbers $\mathbb{N}$ outside the base field of $\mathbb{Z}_2$ just like, say, the Hamming distance function $d_H(T, S) = |T + S|$ on vector spaces over $\mathbb{Z}_2$ in coding theory. [16] Thus the "size of overlap" bra $\langle T|_U : \wp(U) \to \mathbb{N}$ is not to be confused with the dual ("parity of overlap") functional $\varphi_T = \sum_{u_j \in T} \varphi_{u_j} : \wp(U) \to \mathbb{Z}_2$ where $\varphi_{u_j}(\{u_k\}) = \delta_{jk}$ for $U = \{u_1, ..., u_n\}$.

[6] One possible misinterpretation of QM/Sets is to misinterpret the transporting method as an embedding $\mathbb{Z}_2^n \to \mathbb{C}^n$ defined by $\{u_j\} \longmapsto |u_j\rangle$ using a basis for each space. But such an embedding from a vector space over a field of finite characteristic to a vector space of characteristic zero cannot be linear. The repeated sum of a nonzero element in the domain space will eventually be 0 but its repeated nonzero image in the codomain space can never be 0. Indeed in QM/Sets, the brackets $\langle T|_U S\rangle = |T \cap S|$ for $T, T', S \subseteq U$ should be thought of *only* as a measure of the overlap since they are not even linear, e.g., $\langle T + T'|_U S\rangle \neq \langle T|_U S\rangle + \langle T'|_U S\rangle$ whenever $T \cap T' \neq \emptyset$.

[7] The term "$\{u_j\} \cap S'$" is not even defined in general since it is the intersection of subsets $\{u_j\} \subseteq U$ and $S' \subseteq U'$ of two different universe sets U and U'.

[8] We use the double-line notation $\|S\|_U$ for the U-norm of a set to distinguish it from the single-line notation $|S|$ for the cardinality of a set. We also use the double-line notation $\|\psi\|$ for the norm in QM although sometimes the single line notation $|\psi|$ is used elsewhere.

$$f \restriction S = rS \text{ holds} \equiv_{df} f \text{ is constant on the subset } S \text{ with the value } r.$$

This is the QM/sets-version of an *eigenvalue equation* for arbitrary functions on a set $f : U \to \mathbb{R}$. Whenever S satisfies $f \restriction S = rS$ for some r (i.e., is a "level set" of f), then S is said to be an *eigenvector* in the vector space $\wp(U)$ of the numerical attribute $f : U \to \mathbb{R}$, and $r \in \mathbb{R}$ is the associated *eigenvalue*. Each eigenvalue r determines as usual an *eigenspace* $\wp\left(f^{-1}(r)\right)$ of its eigenvectors which is a subspace of the vector space $\wp(U)$. The whole space $\wp(U)$ can be expressed as usual as the direct sum of the eigenspaces: $\wp(U) = \sum_{r \in f(U)} \oplus \wp\left(f^{-1}(r)\right)$. Moreover, for distinct eigenvalues $r \neq r'$, any corresponding eigenvectors $S \in \wp\left(f^{-1}(r)\right)$ and $T \in \wp\left(f^{-1}(r')\right)$ are *orthogonal* in the sense that $\langle T|_U S\rangle = 0$. In general, for vectors $S, T \in \wp(U)$, orthogonality means zero overlap, i.e., disjointness.

The characteristic function $\chi_S : U \to \mathbb{R}$ for $S \subseteq U$ has the eigenvalues of 0 and 1 so it is a numerical attribute that *can* be "internalized" as a linear operator $S \cap () : \wp(U) \to \wp(U)$. Hence in this case, the "eigenvalue equation" $f \restriction T = rT$ for $f = \chi_S$ becomes an actual eigenvalue equation $S \cap T = rT$ for a linear[9] operator $S \cap ()$ with the resulting eigenvalues of 1 and 0, and with the resulting eigenspaces $\wp(S)$ and $\wp(S^c)$ (where S^c is the complement of S) agreeing with those "eigenvalues" and "eigenspaces" defined above for an arbitrary numerical attribute $f : U \to \mathbb{R}$.

The characteristic attributes $\chi_S : U \to \mathbb{R}$ are characterized by the property that their value-wise product, i.e., $(\chi_S \bullet \chi_S)(u) = \chi_S(u)\,\chi_S(u)$, is equal to the attribute value $\chi_S(u)$, and that is reflected in the idempotency of the corresponding operators:

$$\wp(U) \xrightarrow{S\cap()} \wp(U) \xrightarrow{S\cap()} \wp(U) = \wp(U) \xrightarrow{S\cap()} \wp(U).$$

Thus the operators $S \cap ()$ corresponding to the characteristic attributes χ_S are *projection operators*.[10]

The (maximal) eigenvectors $f^{-1}(r)$ for f, with r in the *image* or *spectrum* $f(U) \subseteq \mathbb{R}$, span the set U, i.e., $U = \sum_{r \in f(U)} f^{-1}(r)$. Hence the attribute $f : U \to \mathbb{R}$ has a spectral decomposition in terms of its (projection-defining) characteristic functions:

$$f = \textstyle\sum_{r \in f(U)} r\chi_{f^{-1}(r)} : U \to \mathbb{R}$$
Spectral decomposition of set attribute $f : U \to \mathbb{R}$

which is the QM/sets-version of the spectral decomposition $L = \sum_\lambda \lambda P_\lambda$ of a Hermitian operator L in terms of the projection operators P_λ for its eigenvalues λ.

3.4 Completeness and orthogonality of projection operators

For any vector $S \in \wp(U)$, the operator $S \cap () : \wp(U) \to \wp(U)$ is the linear projection operator to the subspace $\wp(S) \subseteq \wp(U)$. The usual completeness and orthogonality conditions on projection operators P_λ to the eigenspaces of an observable-operator have QM/sets-versions for numerical attributes $f : U \to \mathbb{R}$:

1. completeness: $\sum_\lambda P_\lambda = I : V \to V$ in QM has the QM/sets-version:

$$\textstyle\sum_r f^{-1}(r) \cap () = I : \wp(U) \to \wp(U), \text{ and}$$

2. orthogonality: for $\lambda \neq \mu$, $V \xrightarrow{P_\mu} V \xrightarrow{P_\lambda} V = V \xrightarrow{0} V$ (where 0 is the zero operator) has the QM/sets-version: for $r \neq r'$,

[9] It should be noted that the projection operator $S \cap () : \wp(U) \to \wp(U)$ is not only idempotent but linear, i.e., $(S \cap T_1) + (S \cap T_2) = S \cap (T_1 + T_2)$. Indeed, this is the distributive law when $\wp(U)$ is interpreted as a Boolean ring with intersection as multiplication.

[10] In order for general real-valued attributes to be internalized as linear operators, in the way that characteristic functions χ_S were internalized as projection operators $S \cap ()$, the base field would have to be strengthened to $\mathbb{C}$ and that would take us, *mutatis mutandis*, from the probability calculus of QM/sets to that of full QM.

$$\wp(U) \overset{f^{-1}(r')\cap()}{\longrightarrow} \wp(U) \overset{f^{-1}(r)\cap()}{\longrightarrow} \wp(U) = \wp(U) \overset{0}{\longrightarrow} \wp(U).$$

Note that in spite of the lack of an inner product, the orthogonality of projection operators $S \cap ()$ is perfectly well-defined in QM/sets where it boils down to the disjointness of subsets, i.e., the cardinality of subsets' overlap (instead of their inner product) being 0.

3.5 The Born Rule for measurement in QM and QM/sets

An orthogonal decomposition of a finite set U is just a partition $\pi = \{B, ...\}$ of U since the blocks $B, B', ...$ are orthogonal (i.e., disjoint) and their sum is U. Given such an orthogonal decomposition of U, we have the:

$$\|U\|_U^2 = \sum_{B\in\pi} \|B\|_U^2$$
Pythagorean Theorem
for orthogonal decompositions of sets.

An old question is: "why the squaring of amplitudes in the Born rule of QM?" A superposition state between certain definite orthogonal alternatives A and B, where the latter are represented by vectors $\overrightarrow{A}$ and $\overrightarrow{B}$, is represented by the vector sum $\overrightarrow{C} = \overrightarrow{A} + \overrightarrow{B}$. But what is the "strength," "intensity," or relative importance of the vectors $\overrightarrow{A}$ and $\overrightarrow{B}$ in the vector sum $\overrightarrow{C}$? That question requires a *scalar* measure of strength or intensity. The magnitude or "length" given by the norm $\|\|$ does not answer the question since $\left\|\overrightarrow{A}\right\| + \left\|\overrightarrow{B}\right\| \neq \left\|\overrightarrow{C}\right\|$. But the Pythagorean Theorem shows that the norm-squared gives the scalar measure of "intensity" that answers the question: $\left\|\overrightarrow{A}\right\|^2 + \left\|\overrightarrow{B}\right\|^2 = \left\|\overrightarrow{C}\right\|^2$ in vector spaces over $\mathbb{Z}_2$ or over $\mathbb{C}$. And when the superposition state is reduced by a measurement, then the *probability* that the indefinite state will reduce to one of the definite alternatives is given by that relative scalar measure of the eigen-alternative's "strength" or "intensity" in the indefinite state–and that is the Born Rule. In a slogan, Born is the off-spring of Pythagoras.

Given an observable-operator L in ordinary QM/$\mathbb{C}$ and a numerical attribute in QM/sets, the corresponding Pythagorean Theorems for the complete sets of orthogonal projection operators are:

$$\|\psi\|^2 = \sum_\lambda \|P_\lambda(\psi)\|^2 \text{ and}$$
$$\|S\|_U^2 = \sum_r \left\|f^{-1}(r) \cap S\right\|_U^2 = \sum_r \left|f^{-1}(r) \cap S\right| = |S|.$$

Normalizing gives:

$$\sum_\lambda \frac{\|P_\lambda(\psi)\|^2}{\|\psi\|^2} = 1 \text{ and}$$
$$\sum_r \frac{\left\|f^{-1}(r)\cap S\right\|_U^2}{\|S\|_U^2} = \sum_r \frac{\left|f^{-1}(r)\cap S\right|}{|S|} = 1$$

Here $\frac{\|P_\lambda(\psi)\|^2}{\|\psi\|^2}$ is the "mysterious" quantum probability of getting λ in an L-measurement of ψ, while $\frac{\left\|f^{-1}(r)\cap S\right\|_U^2}{\|S\|_U^2} = \frac{\left|f^{-1}(r)\cap S\right|}{|S|}$ has the rather unmysterious interpretation in the pedagogical model, QM/sets, as the probability $\Pr(r|S)$ of the numerical attribute $f : U \to \mathbb{R}$ having the eigenvalue r when "measuring" $S \in \wp(U)$. Thus the QM/sets-version of the Born Rule is the perfectly ordinary Laplace-Boole rule for the conditional probability $\Pr(r|S) = \frac{\left|f^{-1}(r)\cap S\right|}{|S|}$, that given $S \subseteq U$, a random variable $f : U \to \mathbb{R}$ takes the value r.

In QM/sets, when the indefinite state S is being "measured" using the observable f where the probability $\Pr(r|S)$ of getting the eigenvalue r is $\frac{\left\|f^{-1}(r)\cap S\right\|_U^2}{\|S\|_U^2} = \frac{\left|f^{-1}(r)\cap S\right|}{|S|}$, the "damned quantum jump" (Schrödinger) goes from S by the projection operator $f^{-1}(r) \cap ()$ to the projected resultant

state $f^{-1}(r) \cap S$ which is in the eigenspace $\wp\left(f^{-1}(r)\right)$ for that eigenvalue r. The state resulting from the measurement represents a more-definite state $f^{-1}(r) \cap S$ that now has the definite f-value of r–so a second measurement would yield the same eigenvalue r with probability:

$$\Pr\left(r|f^{-1}(r) \cap S\right) = \frac{\left|f^{-1}(r) \cap \left[f^{-1}(r) \cap S\right]\right|}{|f^{-1}(r) \cap S|} = \frac{\left|f^{-1}(r) \cap S\right|}{|f^{-1}(r) \cap S|} = 1$$

and the same resulting vector $f^{-1}(r) \cap \left[f^{-1}(r) \cap S\right] = f^{-1}(r) \cap S$ using the idempotency of the projection operators.

Hence the treatment of measurement in QM/sets is *all* analogous to the treatment of measurement in standard Dirac-von-Neumann QM.

3.6 Summary of QM/sets and QM

The QM/set-versions of the corresponding QM notions are summarized in the following table for the finite U-basis of the $\mathbb{Z}_2$-vector space $\wp(U)$ and for an orthonormal basis $\{|v_i\rangle\}$ of a finite dimensional Hilbert space V.

QM/sets over $\mathbb{Z}_2$	Standard QM over $\mathbb{C}$
Projections: $S \cap (): \wp(U) \to \wp(U)$	$P : V \to V$ where $P^2 = P$
Spectral Decomposition.: $f = \sum_r r\chi_{f^{-1}(r)}$	$L = \sum_\lambda \lambda P_\lambda$
Completeness.: $\sum_r f^{-1}(r) \cap () = I$	$\sum_\lambda P_\lambda = I$
Orthog.: $r \neq r'$, $\left[f^{-1}(r) \cap ()\right]\left[f^{-1}(r') \cap ()\right] = \emptyset \cap ()$	$\lambda \neq \mu$, $P_\lambda P_\mu = 0$
Brackets: $\langle S|_U T\rangle = \|S \cap T\|$ = overlap of $S, T \subseteq U$	$\langle\psi|\varphi\rangle$ = overlap of ψ and φ
Norm: $\|S\|_U = \sqrt{\langle S|_U S\rangle} = \sqrt{|S|}$ where $S \subseteq U$	$\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$
Pythagoras: $\|S\|_U^2 = \sum_r \left\|f^{-1}(r) \cap S\right\|_U^2$	$\|\psi\|^2 = \sum_\lambda \|P_\lambda(\psi)\|^2$
Normalized: $\sum_r \frac{\left\|f^{-1}(r) \cap S\right\|_U^2}{\|S\|_U^2} = \sum_r \frac{\left\|f^{-1}(r) \cap S\right\|}{\|S\|} = 1$	$\sum_\lambda \frac{\|P_\lambda(\psi)\|^2}{\|\psi\|^2} = 1$
Born rule: $\Pr(r\|S) = \frac{\left\|f^{-1}(r) \cap S\right\|_U^2}{\|S\|_U^2} = \frac{\left\|f^{-1}(r) \cap S\right\|}{\|S\|}$	$\Pr(\lambda\|\psi) = \frac{\|P_\lambda(\psi)\|^2}{\|\psi\|^2}$

Probability calculus for QM/sets over $\mathbb{Z}_2$ and for standard QM over $\mathbb{C}$

4 Measurement in QM/sets

4.1 Measurement, Partitions, and Distinctions

In QM/sets, numerical attributes $f : U \to \mathbb{R}$ can be considered as random variables on a set of equiprobable states $\{u\} \subseteq U$. The inverse images of attributes (or random variables) define set partitions $\left\{f^{-1}\right\} = \left\{f^{-1}(r)\right\}_{r \in f(U)}$ on the set U. Considered abstractly, the partitions on a set U are partially ordered by refinement where a partition $\pi = \{B, ...\}$ *refines* a partition $\sigma = \{C, ...\}$, written $\sigma \preceq \pi$, if for any block $B \in \pi$, there is a block $C \in \sigma$ such that $B \subseteq C$. The principal logical operation needed here is the *partition join* where the join $\pi \vee \sigma$ is the partition whose blocks are the (non-empty) intersections $B \cap C$ for $B \in \pi$ and $C \in \sigma$.

Each partition π can be represented as a binary relation $\mathrm{dit}(\pi) \subseteq U \times U$ on U where the ordered pairs (u, u') in $\mathrm{dit}(\pi)$ are the *distinctions* or *dits* of π in the sense that u and u' are in distinct blocks of π. These *dit sets* $\mathrm{dit}(\pi)$ as binary relations might be called *partition relations* which are also called "apartness relations" in computer science. An ordered pair (u, u') is an *indistinction* or *indit* of π if u and u' are in the same block of π. The set of indits, $\mathrm{indit}(\pi)$, as a binary relation is just the equivalence relation associated with the partition π, the complement of the dit set $\mathrm{dit}(\pi)$ in $U \times U$.

In the category-theoretic duality between *sub*-sets (which are the subject matter of Boole's subset logic, the latter being usually mis-specified as the special case of "propositional" logic) and

quotient-sets or partitions ([9] or [10]), the *elements* of a subset and the *distinctions* of a partition are corresponding concepts.[11]

The partial ordering of subsets in the Boolean lattice $\wp(U)$ is the inclusion of elements, and the refinement partial ordering of partitions in the partition lattice $\prod(U)$ is just the inclusion of distinctions, i.e., $\sigma \preceq \pi$ iff $\mathrm{dit}(\sigma) \subseteq \mathrm{dit}(\pi)$. The top of the Boolean lattice is the subset U of all possible elements and the top of the partition lattice is the *discrete partition* $\mathbf{1} = \{\{u\}\}_{u \in U}$ of singletons which makes all possible distinctions: $\mathrm{dit}(\mathbf{1}) = U \times U - \Delta$ (where $\Delta = \{(u,u) : u \in U\}$ is the diagonal). The bottom of the Boolean lattice is the empty set $\emptyset$ of no elements and the bottom of the lattice of partitions is the *indiscrete partition* (or *blob*) $\mathbf{0} = \{U\}$ which has no distinctions.

The two lattices can be illustrated in the case of $U = \{a, b, c\}$.

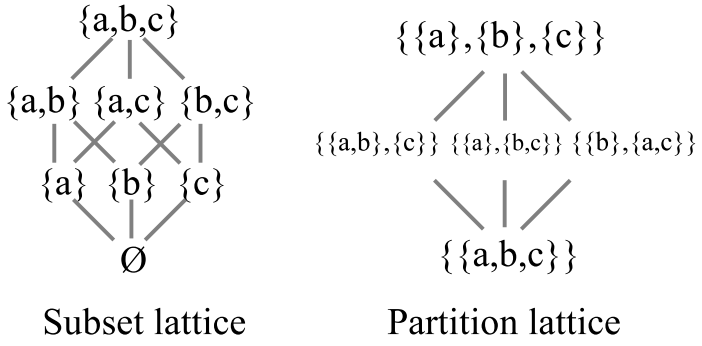


Figure 1: Subset and partition lattices

In the correspondences between QM/sets and QM, a block S in a partition on U [i.e., a vector $S \in \wp(U)$] corresponds to *pure* state in QM, and a partition $\pi = \{B, ...\}$ on U is the *mixed state* of orthogonal pure states B. In QM, a measurement makes distinctions, i.e., makes alternatives distinguishable, and that turns a pure state into a mixture of probabilistic outcomes. A measurement in QM/sets is the distinction-creating process of turning a pure state $S \in \wp(U)$ into a mixed state partition $\left\{f^{-1}(r) \cap S\right\}_{r \in f(U)}$ on S. The distinction-creating process of measurement in QM/sets is the action on S of the inverse-image partition $\left\{f^{-1}(r)\right\}_{r \in f(U)}$ in the join $\{S, S^c\} \vee \left\{f^{-1}(r)\right\}$ with the partition $\{S, S^c\}$, so that action on S is:

$$S \longrightarrow \left\{f^{-1}(r) \cap S\right\}_{r \in f(U)}$$

Action on the pure state S of an f-measurement-join to give mixed state $\left\{f^{-1}(r) \cap S\right\}_{r \in f(U)}$ on S.

The nonempty states $\left\{f^{-1}(r) \cap S\right\}_{r \in f(U)}$ are all possible or "potential" but the actual indefinite state S turns into one of the definite states with the probabilities given by the probability calculus: $\Pr(r|S) = \frac{\left\|f^{-1}(r) \cap S\right\|_U^2}{\|S\|_U^2} = \frac{\left|f^{-1}(r) \cap S\right|}{|S|}$. Since the reduction of the state S to the state $f^{-1}(r) \cap S$ is mathematically described by applying the projection operator $f^{-1}(r) \cap ()$, it is called a *projective* measurement.

The pedagogical model, QM/sets, could be seen as a development of some of the hints in Hermann Weyl's expository writings about quantum mechanics. He called a partition a "grating" or "sieve"[12], and then considered *both* set partitions and vector space partitions (direct sum decompositions) as the respective types of gratings.[22, pp. 255-257] He started with a numerical attribute

[11] Boole has been included along with Laplace in the name of classical finite probability theory since he developed it as the normalized counting measure on the elements of the subsets of his logic. Applying the same mathematical move to the dual logic of partitions results in developing the notion of *logical entropy* $h(\pi)$ of a partition π as the normalized counting measure on the dit set $\mathrm{dit}(\pi)$, i.e., $h(\pi) = \frac{|\mathrm{dit}(\pi)|}{|U \times U|}$. [8]

[12] Arthur Eddington made a very early use of the sieve idea:

on a set, e.g., $f : U \to \mathbb{R}$, which defined the set partition or "grating" [22, p. 255] with blocks having the same attribute-value, e.g., $\left\{f^{-1}(r)\right\}_{r \in f(U)}$. Then he moved to the QM case where the universe set, e.g., $U = \{u_1, ..., u_n\}$, or "aggregate of n states has to be replaced by an n-dimensional Euclidean vector space" [22, p. 256]. The appropriate notion of a vector space partition or "grating" is a "splitting of the total vector space into mutually orthogonal subspaces" so that "each vector $\overrightarrow{x}$ splits into r component vectors lying in the several subspaces" [22, p. 256], i.e., a direct sum decomposition of the space. After referring to a partition as a "grating" or "sieve," Weyl notes that "Measurement means application of a sieve or grating" [22, p. 259], e.g., in QM/sets, the application (i.e., join) of the set-grating or partition $\left\{f^{-1}(r)\right\}_{r \in f(U)}$ to the pure state $\{S\}$ to give the mixed state $\left\{f^{-1}(r) \cap S\right\}_{r \in f(U)}$.

For some visual imagery of measurement, we might think of the grating as a series of regular-polygonal-shaped holes that might shape an indefinite blob of dough. In a measurement, the blob of dough falls through one of the polygonal holes in the grating with equal probability and then takes on that shape.

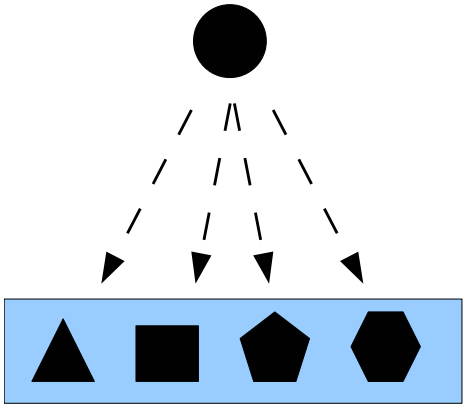


Figure 2: Measurement as randomly giving an indefinite blob of dough a definite polygonal shape.

4.2 Measurement in QM/Sets

In the simple example illustrated below, we start at the one block or state of the indiscrete partition or blob which is $\{a, b, c\}$. A measurement uses some attribute that defines an inverse-image partition on $U = \{a, b, c\}$. In the case at hand, there are "essentially" four possible attributes that could be used to "measure" the state $\{a, b, c\}$ (since there are four partitions that refine the indiscrete partition).

For an example of a degenerate measurement, we choose an attribute with a non-discrete inverse-image partition such as the partition $\pi = \{\{a\}, \{b, c\}\}$ which determines a DSD $\{\wp(\{a\}), \wp(\{b, c\})\}$. Hence the attribute could just be the characteristic function $\chi_{\{b,c\}}$ with the two eigenspaces $\wp(\{a\})$ and $\wp(\{b, c\})$ and the two eigenvalues 0 and 1 respectively. Since the eigenspace $\wp\left(\chi^{-1}_{\{b,c\}}(1)\right) = \wp(\{b, c\})$ is not one dimensional, the eigenvalue of 1 is a QM/Sets-version of a *degenerate* eigenvalue. This attribute $\chi_{\{b,c\}}$ has four (non-zero) eigenvectors:

$$\chi_{\{b,c\}} \upharpoonright \{b, c\} = 1\{b, c\},\ \chi_{\{b,c\}} \upharpoonright \{b\} = 1\{b\},\ \chi_{\{b,c\}} \upharpoonright \{c\} = 1\{c\}, \text{ and } \chi_{\{b,c\}} \upharpoonright \{a\} = 0\{a\}.$$

The "measuring apparatus" makes distinctions by joining the attribute's inverse-image partition

> In Einstein's theory of relativity the observer is a man who sets out in quest of truth armed with a measuring-rod. In quantum theory he sets out armed with a sieve.[7, p. 267]

This passage was quoted by Weyl [22, p. 255] in his treatment of gratings.

$$\chi^{-1}_{\{b,c\}} = \left\{\chi^{-1}_{\{b,c\}}(1), \chi^{-1}_{\{b,c\}}(0)\right\} = \{\{b,c\},\{a\}\}$$

with the pure state representing the indefinite entity $U = \{a,b,c\}$. The action on the pure state is:

$$U \to \{U\} \vee \chi^{-1}_{\{b,c\}} = \chi^{-1}_{\{b,c\}} = \{\{b,c\},\{a\}\}.$$

The measurement of that attribute returns one of the eigenvalues with the probabilities:

$$\Pr(0|U) = \frac{|\{a\}\cap\{a,b,c\}|}{|\{a,b,c\}|} = \frac{1}{3} \text{ and } \Pr(1|U) = \frac{|\{b,c\}\cap\{a,b,c\}|}{|\{a,b,c\}|} = \frac{2}{3}.$$

Suppose it returns the eigenvalue 1. Then the indefinite entity $\{a,b,c\}$ reduces to the projected eigenstate $\chi^{-1}_{\{b,c\}}(1) \cap \{a,b,c\} = \{b,c\}$ for that eigenvalue [4, p. 221].

Since this is a degenerate result (i.e., the eigenspace $\wp\left(\chi^{-1}_{\{b,c\}}(1)\right) = \wp(\{b,c\})$ doesn't have dimension one), another measurement is needed to make more distinctions. Measurements by attributes, such as $\chi_{\{a,b\}}$ or $\chi_{\{a,c\}}$, that give either of the other two partitions, $\{\{a,b\},\{c\}\}$ or $\{\{b\},\{a,c\}\}$ as inverse images, would suffice to distinguish $\{b,c\}$ into $\{b\}$ or $\{c\}$. Then either attribute together with the attribute $\chi_{\{b,c\}}$ would form a *Complete Set of Compatible Attributes* or CSCA (i.e., the QM/Sets-version of Dirac's Complete Set of Commuting Operators or CSCO), where *complete* means that the join of the attributes' inverse-image partitions gives the discrete partition and where *compatible* means that all the attributes can be taken as defined on the same set of (simultaneous) basis eigenvectors, e.g., the U-basis.

Taking, for example, the other attribute as $\chi_{\{a,b\}}$, the join of the two attributes' partitions is discrete:

$$\chi^{-1}_{\{b,c\}} \vee \chi^{-1}_{\{a,b\}} = \{\{a\},\{b,c\}\} \vee \{\{a,b\},\{c\}\} = \{\{a\},\{b\},\{c\}\} = \mathbf{1}.$$

Hence all the eigenstate singletons can be characterized by the ordered pairs of the eigenvalues of these two attributes: $\{a\} = |0,1\rangle$, $\{b\} = |1,1\rangle$, and $\{c\} = |1,0\rangle$ (using Dirac's ket-notation to give the ordered pairs and listing the eigenvalues of $\chi_{\{b,c\}}$ first on the left).

The second projective measurement of the indefinite entity $\{b,c\}$ using the attribute $\chi_{\{a,b\}}$ with the inverse-image partition $\chi^{-1}_{\{a,b\}} = \{\{a,b\},\{c\}\}$ would have the pure-to-mixed state action:

$$\{b,c\} \to \left\{\{b,c\} \cap \chi_{\{a,b\}}(1), \{b,c\} \cap \chi_{\{a,b\}}(0)\right\} = \{\{b\},\{c\}\}.$$

The distinction-making measurement would cause the indefinite entity $\{b,c\}$ to turn into one of the definite entities of $\{b\}$ or $\{c\}$ with the probabilities:

$$\Pr(1|\{b,c\}) = \frac{|\{a,b\}\cap\{b,c\}|}{|\{b,c\}|} = \frac{1}{2} \text{ and } \Pr(0|\{b,c\}) = \frac{|\{c\}\cap\{b,c\}|}{|\{b,c\}|} = \frac{1}{2}.$$

If the measured eigenvalue is 0, then the state $\{b,c\}$ projects to $\chi^{-1}_{\{a,b\}}(0) \cap \{b,c\} = \{c\}$ as pictured below.

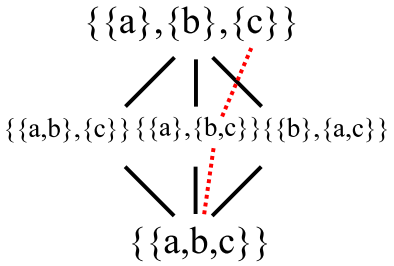


Figure 3: Degenerate measurement

The two projective measurements of $\{a,b,c\}$ using the complete set of compatible (e.g., both defined on U) attributes $\chi_{\{b,c\}}$ and $\chi_{\{a,b\}}$ produced the respective eigenvalues 1 and 0 so the resulting eigenstate was characterized by the eigenket $|1,0\rangle = \{c\}$.

Again, this is all analogous to standard Dirac-von-Neumann quantum mechanics.

5 Density matrices and measurement in QM/sets

The previous treatment of the role of partitions in measurement can be restated using density matrices over the reals. Given a partition $\pi = \{B, ...\}$ on $U = \{u_1, ..., u_n\}$, the blocks $B \in \pi$ can be thought of as (nonoverlapping or "orthogonal") "pure states" where the "state" B occurs with the probability $p_B = \frac{|B|}{|U|}$. Then we can transport the usual procedure for forming the density matrix $\rho(\pi)$ for the "orthogonal pure states" B with the probabilities p_B. The "pure state" B normalized in the reals to length 1 is represented by the column vector $|B\rangle_1 = \frac{1}{\sqrt{|B|}}[\chi_B(u_1), ..., \chi_B(u_n)]^t$(where $[\,]^t$ indicates the transpose). Then the *density matrix* $\rho(B)$ *for the pure state* $B \subseteq U$ is then (calculating in the reals):

$$\rho(B) = |B\rangle_1 (|B\rangle_1)^t = \frac{1}{|B|} \begin{bmatrix} \chi_B(u_1) \\ \chi_B(u_2) \\ \vdots \\ \chi_B(u_n) \end{bmatrix} [\chi_B(u_1), ..., \chi_B(u_n)]$$

$$= \frac{1}{|B|} \begin{bmatrix} \chi_B(u_1) & \chi_B(u_1)\chi_B(u_2) & \cdots & \chi_B(u_1)\chi_B(u_n) \\ \chi_B(u_2)\chi_B(u_1) & \chi_B(u_2) & \cdots & \chi_B(u_2)\chi_B(u_n) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_B(u_n)\chi_B(u_1) & \chi_B(u_n)\chi_B(u_2) & \cdots & \chi_B(u_n) \end{bmatrix}.$$

For instance if $U = \{u_1, u_2, u_3\}$, then for the blocks in the partition $\pi = \{\{u_1, u_2\}, \{u_3\}\}$:

$$\rho(\{u_1, u_2\}) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \rho(\{u_3\}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then the "mixed state" *density matrix* $\rho(\pi)$ *of the partition* π is the weighted sum:

$$\rho(\pi) = \textstyle\sum_{B \in \pi} p_B \rho(B).$$

In the example, this is:

$$\rho(\pi) = \frac{2}{3} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.$$

In partition logic [10], given a set partition $\pi = \{B, B', ...\}$ on a universe set U, an ordered pair $(u, u') \in U \times U$ is called a *distinction* or *dit* of π if the elements are in different blocks of π, and the set of all distinctions is the *dit set* $\mathrm{dit}(\pi)$. An ordered pair (u, u') is called an *indistinction* or *indit* of π if the two elements are in the same block of π, and the set of all indistinctions is the indit set $\mathrm{indit}(\pi)$. A partition π has an associated binary equivalence relation which is its indit set $\mathrm{indit}(\pi) \subseteq U \times U$, and an associated partition relation or apartness relation which is the complementary dit set $\mathrm{dit}(\pi) = U \times U - \mathrm{indit}(\pi)$. The density matrix $\rho(\pi)$ of the partition can then be directly interpreted in terms of its indit set:

$$\rho_{jk}(\pi) = \begin{cases} \frac{1}{|U|} \text{ if } (u_j, u_k) \in \text{indit}(\pi) \\ 0 \text{ if } (u_j, u_k) \notin \text{indit}(\pi) \end{cases}.$$

All the entries are real "amplitudes" whose sum of squares are the two-draw probabilities of drawing a pair of elements from U (with replacement) that is an indistinction of π. Like in the full quantum case, the non-zero entries of the density matrix $\rho_{jk}(\pi) = \sqrt{\frac{1}{|U|}\frac{1}{|U|}} = \frac{1}{|U|}$ are the "coherences" [4, p. 302] which indicate that u_j and u_k "cohere" together in a block or "pure state" of the partition, i.e., for some block $B \in \pi$, $u_j, u_k \in B$. Since the ordered pairs (u_j, u_j) in the diagonal $\Delta \subseteq U \times U$ are always indits of any partition, the diagonal entries in $\rho(\pi)$ are always $\frac{1}{|U|}$.

Combinatorial theory gives a natural way to define the same density matrix $\rho(\pi)$ of a partition π. A binary relation $R \subseteq U \times U$ on $U = \{u_1, ..., u_n\}$ can be represented by an $n \times n$ *incidence matrix* $I(R)$ where

$$I(R)_{jk} = \begin{cases} 1 \text{ if } (u_j, u_k) \in R \\ 0 \text{ if } (u_j, u_k) \notin R. \end{cases}$$

Taking R as the equivalence relation $\text{indit}(\pi)$ associated with a partition π, the density matrix $\rho(\pi)$ defined above is just the incidence matrix $I(\text{indit}(\pi))$ rescaled to be of trace 1 (i.e., sum of diagonal entries is 1):

$$\rho(\pi) = \frac{1}{|U|} I(\text{indit}(\pi)).$$

If the subsets $T \in \wp(U)$ are represented by the n-ary column vectors $[\chi_T(u_1), ..., \chi_T(u_n)]^t$, then the action of the projection operator $B \cap (): \wp(U) \to \wp(U)$ is represented in the U-basis by the $n \times n$ diagonal matrix P_B where the diagonal entries are:

$$(P_B)_{jj} = \begin{cases} 1 \text{ if } u_j \in B \\ 0 \text{ if } u_j \notin B \end{cases} = \chi_B(u_j)$$

which is idempotent, $P_B^2 = P_B$, and symmetric, $P_B^t = P_B$. For any state $S \in \wp(U)$, the trace (sum of diagonal entries) of $P_B \rho(S)$ is:

$$\text{tr}[P_B \rho(S)] = \frac{1}{|S|} \sum_{j=1}^{n} \chi_S(u_j) \chi_B(u_j) = \frac{|B \cap S|}{|S|} = \Pr(B|S)$$

so given $f: U \to \mathbb{R}$,

$$\Pr(r|S) = \frac{\left|f^{-1}(r) \cap S\right|}{|S|} = \text{tr}\left[P_{f^{-1}(r)} \rho(S)\right]$$

This is the QM/Sets version of the usual result: $\Pr(\lambda|\psi) = \frac{\|P_\lambda(\psi)\|^2}{\|\psi\|^2} = \text{tr}[P_\lambda \rho(\psi)]$.

Given a state S, the measurement by the f-attribute DSD $\left\{\wp\left(f^{-1}(r)\right)\right\}_{r \in f(U)}$ projects S to the state $f^{-1}(r) \cap S$ with the probability $\text{tr}[P_{f^{-1}(r)} \rho(S)] = \frac{|f^{-1}(r) \cap S|}{|S|} = \Pr(r|S)$. We need to convert this into the language of density matrices. Starting with the pure state S as a normalized column vector $|S\rangle_1$, the subset $f^{-1}(r) \cap S$ resulting from that projection is the column vector $P_{f^{-1}(r)}|S\rangle$. To calculate the corresponding density matrix we must first normalize the column vector $P_{f^{-1}(r)}|S\rangle$ by dividing through by $\sqrt{|f^{-1}(r) \cap S|}$ (where nonzero). But the normalizing factor to compute $\rho(S)$ was $\sqrt{|S|}$, i.e., $|S\rangle_1 = \frac{1}{\sqrt{|S|}}|S\rangle$. Since $\text{tr}\left[P_{f^{-1}(r)} \rho(S)\right] = \frac{\left|f^{-1}(r) \cap S\right|}{|S|}$, the normalized version of $P_{f^{-1}(r)}|S\rangle$ is:

$$\frac{1}{\sqrt{|f^{-1}(r) \cap S|}} P_{f^{-1}(r)}|S\rangle = \frac{1}{\sqrt{|f^{-1}(r) \cap S|}} P_{f^{-1}(r)} \sqrt{|S|}\,|S\rangle_1 = \frac{1}{\sqrt{\text{tr}\left[P_{f^{-1}(r)} \rho(S)\right]}} P_{f^{-1}(r)}|S\rangle_1.$$

Hence the density matrix corresponding to the projected state $P_{f^{-1}(r)}\left|S\right\rangle$ is:

$$\frac{1}{\mathrm{tr}\left[P_{f^{-1}(r)}\rho(S)\right]}\left(P_{f^{-1}(r)}\left|S\right\rangle_1\right)\left(P_{f^{-1}(r)}\left|S\right\rangle_1\right)^t$$
$$=\frac{1}{\mathrm{tr}\left[P_{f^{-1}(r)}\rho(S)\right]}P_{f^{-1}(r)}\left|S\right\rangle_1\left(\left|S\right\rangle_1\right)^t\left(P_{f^{-1}(r)}\right)^t=\frac{P_{f^{-1}(r)}\rho(S)P_{f^{-1}(r)}}{\mathrm{tr}\left[P_{f^{-1}(r)}\rho(S)\right]}.$$

This might be illustrated by using the degenerate measurement where $f=\chi_{\{a,b\}}$ and $S=\{b,c\}$. Then the density matrix is:

$$\rho\left(\{b,c\}\right)=\begin{bmatrix}0&0&0\\0&\frac{1}{2}&\frac{1}{2}\\0&\frac{1}{2}&\frac{1}{2}\end{bmatrix}\text{ and }\chi^{-1}_{\{a,b\}}\left(1\right)=f^{-1}\left(1\right)=\{a,b\}\text{ so }P_{f^{-1}(1)}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}.$$
$$P_{f^{-1}(1)}\rho\left(\{b,c\}\right)=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}0&0&0\\0&\frac{1}{2}&\frac{1}{2}\\0&\frac{1}{2}&\frac{1}{2}\end{bmatrix}=\begin{bmatrix}0&0&0\\0&\frac{1}{2}&\frac{1}{2}\\0&0&0\end{bmatrix}$$
$$P_{f^{-1}(1)}\rho\left(\{b,c\}\right)P_{f^{-1}(1)}=\begin{bmatrix}0&0&0\\0&\frac{1}{2}&\frac{1}{2}\\0&0&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}=\begin{bmatrix}0&0&0\\0&\frac{1}{2}&0\\0&0&0\end{bmatrix}$$

Since $\mathrm{tr}\left[P_{f^{-1}(1)}\rho\left(\{b,c\}\right)\right]=\frac{1}{2}$, the resultant state from that projection is:

$$\frac{P_{f^{-1}(1)}\rho(\{b,c\})P_{f^{-1}(1)}}{\mathrm{tr}\left[P_{f^{-1}(1)}\rho(\{b,c\})\right]}=\frac{1}{1/2}\begin{bmatrix}0&0&0\\0&\frac{1}{2}&0\\0&0&0\end{bmatrix}=\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}$$

with the density matrix $\rho\left(\{b\}\right)$ where $\{b\}=f^{-1}\left(1\right)\cap\{b,c\}=\{a,b\}\cap\{b,c\}$. For the other eigenvalue of 0, we have

$$P_{f^{-1}(0)}\rho\left(\{b,c\}\right)P_{f^{-1}(0)}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0&0&0\\0&\frac{1}{2}&\frac{1}{2}\\0&\frac{1}{2}&\frac{1}{2}\end{bmatrix}\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$$
$$=\begin{bmatrix}0&0&0\\0&0&0\\0&\frac{1}{2}&\frac{1}{2}\end{bmatrix}\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&\frac{1}{2}\end{bmatrix}$$

and $\mathrm{tr}\left[P_{f^{-1}(0)}\rho\left(\{b,c\}\right)\right]=\frac{1}{2}$ so

$$\frac{P_{f^{-1}(0)}\rho(\{b,c\})P_{f^{-1}(0)}}{\mathrm{tr}\left[P_{f^{-1}(0)}\rho(\{b,c\})\right]}=\frac{1}{1/2}\begin{bmatrix}0&0&0\\0&0&0\\0&0&\frac{1}{2}\end{bmatrix}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$$

which is the density matrix for the pure state $\{c\}=f^{-1}\left(0\right)\cap\{b,c\}=\{c\}\cap\{b,c\}$.

The final formula for the post-measurement mixed state $\hat{\rho}\left(S\right)$ would weigh the projected states by their probability, so we have:

$$\hat{\rho}\left(S\right)=\sum_{r\in f(U)}\Pr\left(r|S\right)\frac{P_{f^{-1}(r)}\rho(S)P_{f^{-1}(r)}}{\mathrm{tr}\left[P_{f^{-1}(r)}\rho(S)\right]}$$
$$=\sum_{r\in f(U)}\mathrm{tr}\left[P_{f^{-1}(r)}\rho\left(S\right)\right]\frac{P_{f^{-1}(r)}\rho(S)P_{f^{-1}(r)}}{\mathrm{tr}\left[P_{f^{-1}(r)}\rho(S)\right]}=\sum_{r\in f(U)}P_{f^{-1}(r)}\rho\left(S\right)P_{f^{-1}(r)}.$$

Thus the action of the measurement is:

$$\rho\left(S\right)\longrightarrow\hat{\rho}\left(S\right)=\sum_{r\in f(U)}P_{f^{-1}(r)}\rho\left(S\right)P_{f^{-1}(r)}$$
Measurement of S using f-attribute in density matrix form.

This result is just the "transported" QM/Sets version of the description of measurement in full QM. Consider the projective measurement using a self-adjoint operator F on V with the DSD $\{V_\lambda\}$ of eigenspaces and the projections to the eigenspaces $P_\lambda : V \to V_\lambda$. The measurement of a normalized pure state $|\psi\rangle$ results in the state $P_\lambda |\psi\rangle$ with the probability $p_\lambda = \mathrm{tr}\left[P_\lambda \rho\left(\psi\right)\right] = \Pr\left(\lambda|\psi\right)$ where $\rho\left(\psi\right) = |\psi\rangle\langle\psi|$. The projected resultant state $P_\lambda |\psi\rangle$ has the density matrix $\frac{P_\lambda|\psi\rangle\langle\psi|P_\lambda}{\mathrm{tr}[P_\lambda\rho(\psi)]} = \frac{P_\lambda\rho(\psi)P_\lambda}{\mathrm{tr}[P_\lambda\rho(\psi)]}$ so the mixed state describing the probabilistic results of the measurement is [17, p. 101 or p. 515]:

$$\hat{\rho}\left(\psi\right) = \sum_\lambda p_\lambda \frac{P_\lambda\rho(\psi)P_\lambda}{\mathrm{tr}[P_\lambda\rho(\psi)]} = \sum_\lambda \mathrm{tr}\left[P_\lambda\rho\left(\psi\right)\right] \frac{P_\lambda\rho(\psi)P_\lambda}{\mathrm{tr}[P_\lambda\rho(\psi)]} = \sum_\lambda P_\lambda\rho\left(\psi\right)P_\lambda.$$

Thus we see how the density matrix treatment of measurement in QM/Sets

$$\rho\left(S\right) \longrightarrow \hat{\rho}\left(S\right) = \sum_{r\in f(U)} P_{f^{-1}(r)}\rho\left(S\right)P_{f^{-1}(r)}$$

is just a sets-version of the density matrix treatment of projective measurement in standard Dirac-von-Neumann QM:

$$\rho\left(\psi\right) \longmapsto \hat{\rho}\left(\psi\right) = \sum_\lambda P_\lambda\rho\left(\psi\right)P_\lambda.$$

6 Commutators and the Indeterminacy Principle in QM/sets

The only attributes $f : U \to \mathbb{R}$ on some basis set U for $\mathbb{Z}_2^n$ (where $|U| = n$) that can be *internalized* (or "quantized") as linear operators $\mathbb{Z}_2^n \to \mathbb{Z}_2^n$ are the characteristic functions $\chi_S : U \to 2 \subseteq \mathbb{R}$. But the properties of a general attribute $f : U \to \mathbb{R}$ can be analyzed in terms of the projection operators $P_{f^{-1}(r)}$, of the characteristic functions $\chi_{f^{-1}(r)} : U \to 2$. Hence we focus on projection operators $P : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$.

Given two projection operators $P, P' : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$, they may come from characteristic functions $\chi_S : U \to 2$ and $\chi_{S'} : U' \to 2$ for quite different basis sets U and U'. But each projection has eigenvalues $0, 1$ so they define two Direct Sum Decompositions (DSDs) in terms of their eigenspaces (where $V_i = \wp\left(\chi_S^{-1}\left(i\right)\right)$ and similarly for V_i'):

$$V_1 \oplus V_0 = \mathbb{Z}_2^n = V_1' \oplus V_0'.$$

If we think of the eigenspaces as being like the disjoint blocks in two binary set partitions, we can then try to form the *partition join* by considering the pairwise intersections of the eigenspaces and the space K they span:

$$K = (V_1 \cap V_1') \oplus (V_1 \cap V_0') \oplus (V_0 \cap V_1') \oplus (V_0 \cap V_0') \subseteq \mathbb{Z}_2^n.$$
Simultaneous eigenvector space

The non-zero vectors in those intersections $V_i \cap V_j'$ are, by definition, the *simultaneous eigenvectors* of the two operators P and P', so the space K is the space spanned by the simultaneous eigenvectors.

The non-zero subspaces $V_i \cap V_j'$ form a DSD of K which might be called the *proto-join* of the two DSDs $\{V_1, V_0\}$ and $\{V_1', V_0'\}$ ("proto" since K might not be the whole space). If K is the whole space $\mathbb{Z}_2^n$, then that DSD would be called the *join of the DSDs* $\{V_1, V_0\}$ and $\{V_1', V_0'\}$.

What is the subspace K and when is it the whole space $\mathbb{Z}_2^n$?

The *kernel* of a linear operator $L : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ is the subspace of all vectors that are mapped to the zero vector. Given the two linear operators $P, P' : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$, their *commutator* is the operator:

$$[P, P'] = PP' - P'P : \mathbb{Z}_2^n \to \mathbb{Z}_2^n.$$

Theorem: $K = \ker([P, P'])$, i.e., the simultaneous eigenvector space is the kernel of the commutator.

Proof: While we will only use this theorem for the case of projection operators on $\mathbb{Z}_2^n$, it is true for any diagonalizable operators on any finite dimensional vector space. Let $L, M : V \to V$ be two diagonalizable operators on a finite dimensional vector space V and let v be a simultaneous eigenvector of the operators, i.e., $Lv = \lambda v$ and $Mv = \mu v$. Then $[L, M](v) = (LM - ML)(v) = (\lambda\mu - \mu\lambda) v = 0$ so the space K spanned by the simultaneous eigenvectors is contained in the kernel $\ker([L, M])$, i.e., $K \subseteq \ker([L, M])$. Conversely, if we restrict the two operators to the subspace $\ker([L, M])$, then the restricted operators commute on that subspace. Then it is a standard theorem of linear algebra [13, p. 177] that the space is spanned by simultaneous eigenvectors of the two restricted operators. But if a vector is a simultaneous eigenvectors for the two operators restricted to a subspace, they are the same for the operators on the whole space V, so $\ker([L, M]) \subseteq K$. □

Then the operators commute, i.e., $[P, P'] = 0$ (the zero operator) so $\ker([P, P']) = \mathbb{Z}_2^n$ iff the whole space is spanned by the simultaneous eigenvectors of the two operators.

The opposite special case is when the commutator is non-singular so its kernel is the zero space, and we could say the projections are "conjugate incompatible" operators like position and momentum in full QM.

In general, there are three types of compatibility or incompatibility between observables:

- compatible = commutator is the zero operator;
- incompatible = commutator is not the zero operator;
- conjugate = commutator is non-singular (maximal incompatibility).

To simplify notation, subsets like $\{a, b, c\}$ will sometimes be written as abc without the curly brackets and commas.

Example of conjugate operators: In $\mathbb{Z}_2^4$, take the computational U-basis as $\{\{a\}, \{b\}, \{c\}, \{d\}\}$ or $\{a, b, c, d\}$ (without the curly brackets on the subset-vectors) with an operator defined by the U-attribute $f = \chi_{\{a,c\}} : a, c \longmapsto 1$ and $b, d \longmapsto 0$. Let $\hat{U} = \left\{\hat{a}, \hat{b}, \hat{c}, \hat{d}\right\}$ be the "complementary" basis where $\hat{a} = bcd, \hat{b} = acd, \hat{c} = abd$, and $\hat{d} = abc$ [N.B. Those subsets complementary to the singletons only form a basis when U has *even* cardinality]. Consider the operator defined by the $\hat{U}$-attribute $g : \hat{U} \to 2$ where $g = \chi_{\left\{\hat{a},\hat{b}\right\}} : \hat{a}, \hat{b} \longmapsto 1$ and $\hat{c}, \hat{d} \longmapsto 0$. Then the two direct sum decompositions defined by the attributes are:

$$\begin{gathered}\wp\left(f^{-1}(1)\right) \oplus \wp\left(f^{-1}(0)\right) \\ = \{\emptyset, a, c, ac\} \oplus \{\emptyset, b, d, bd\} \\ \text{and} \\ \wp\left(g^{-1}(1)\right) \oplus \wp\left(g^{-1}(0)\right) \\ = \left\{\emptyset, \hat{a}, \hat{b}, \hat{a}\hat{b}\right\} \oplus \left\{\emptyset, \hat{c}, \hat{d}, \hat{c}\hat{d}\right\} \\ = \{\emptyset, bcd, acd, ab\} \oplus \{\emptyset, abd, abc, cd\}.\end{gathered}$$

These two DSD's have no non-zero vectors in common, i.e., $K = \{\emptyset\}$, so the commutator is non-singular.

Let's work through the conversion matrices. The matrix to convert from the $\hat{U}$-basis to the computational U-basis is:

$$C_{U \leftarrow \hat{U}} = \begin{bmatrix} \langle a|_U \hat{a}\rangle & \left\langle a|_U \hat{b}\right\rangle & \langle a|_U \hat{c}\rangle & \left\langle a|_U \hat{d}\right\rangle \\ \langle b|_U \hat{a}\rangle & \left\langle b|_U \hat{b}\right\rangle & \langle b|_U \hat{c}\rangle & \left\langle b|_U \hat{d}\right\rangle \\ \langle c|_U \hat{a}\rangle & \left\langle c|_U \hat{b}\right\rangle & \langle c|_U \hat{c}\rangle & \left\langle c|_U \hat{d}\right\rangle \\ \langle d|_U \hat{a}\rangle & \left\langle d|_U \hat{b}\right\rangle & \langle d|_U \hat{c}\rangle & \left\langle d|_U \hat{d}\right\rangle \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

$$\text{inverse (in rationals) } C_{\hat{U}\leftarrow U}\text{: } \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix} \overset{\text{mod}(2)}{=} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} = C_{\hat{U}\leftarrow U}$$

so the matrix is its own inverse mod(2). Then the matrix in the U-basis for the g-operator $\hat{a}, \hat{b} \to 1; \hat{c}, \hat{d} \to 0$ is:

$$\begin{aligned} & C_{U\leftarrow\hat{U}} P_{\hat{a},\hat{b}} C_{\hat{U}\leftarrow U} \\ &= \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{bmatrix} \overset{\text{mod}(2)}{=} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Then we can compute the commutator in the computational basis:

$$\begin{aligned} & \left[P_{a,c}, P_{\hat{a},\hat{b}}\right] = P_{a,c} P_{\hat{a},\hat{b}} - P_{\hat{a},\hat{b}} P_{a,c} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \overset{\text{mod}(2)}{=} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \sqrt{I} \end{aligned}$$

which is obviously non-singular. Indeed, it is the permutation matrix that makes the interchanges: $a \leftrightarrow d$ and $b \leftrightarrow c$, and thus it is a square root of the identity matrix.

For the conjugate operators like $P_{a,c}$ and $P_{\hat{a},\hat{b}}$, there is no state in which both "observables" have a definite value.

Example of incompatible operators: Replace the operator f in the computational basis by $h : U \to 2$ where $h = \chi_{\{a,b\}} : a, b \longmapsto 1$ and $c, d \longmapsto 0$. Then the two direct sum decompositions of eigenspaces are:

$$\wp\left(h^{-1}(1)\right) \oplus \wp\left(h^{-1}(0)\right) = \{\emptyset, a, b, ab\} \oplus \{\emptyset, c, d, cd\}$$
$$\text{and}$$
$$\wp\left(g^{-1}(1)\right) \oplus \wp\left(g^{-1}(0)\right) = \{\emptyset, bcd, acd, ab\} \oplus \{\emptyset, abd, abc, cd\}.$$

We see that the simultaneous eigenvectors are ab and cd so the kernel of the commutator to be the eigenspace generated by those simultaneous eigenvectors: $\ker\left(\left[P_{a,b}, P_{\hat{a},\hat{b}}\right]\right) = \{\emptyset, ab, cd, abcd\}$. Since this is neither the whole space nor the zero subspace, these two operators are incompatible but not conjugate.

The calculations of that commutator can be carried out for all the sixteen states using the format of a ket table.

U	$\hat{U}$	$h \upharpoonright = P_{a,b}$	$g \upharpoonright = P_{\hat{a},\hat{b}}$	gh	hg	$[g,h]$
$abcd$	$\hat{a}\hat{b}\hat{c}\hat{d}$	ab	$\hat{a}\hat{b} = ab$	ab	ab	0
abc	$\hat{d}$	ab	$\emptyset$	ab	0	ab
abd	$\hat{c}$	ab	$\emptyset$	ab	0	ab
acd	$\hat{b}$	a	$\hat{b} = acd$	acd	a	cd
bcd	$\hat{a}$	b	$\hat{a} = bcd$	bcd	b	cd
ab	$\hat{a}\hat{b}$	ab	$\hat{a}\hat{b} = ab$	ab	ab	0
ac	$\hat{a}\hat{c}$	a	$\hat{a} = bcd$	acd	b	$abcd$
ad	$\hat{a}\hat{d}$	a	$\hat{a} = bcd$	acd	b	$abcd$
bc	$\hat{b}\hat{c}$	b	$\hat{b} = acd$	bcd	a	$abcd$
bd	$\hat{b}\hat{d}$	b	$\hat{b} = acd$	bcd	a	$abcd$
cd	$\hat{c}\hat{d}$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	0
d	$\hat{a}\hat{b}\hat{c}$	$\emptyset$	$\hat{a}\hat{b} = ab$	$\emptyset$	ab	ab
c	$\hat{a}\hat{b}\hat{d}$	$\emptyset$	$\hat{a}\hat{b} = ab$	$\emptyset$	ab	ab
b	$\hat{a}\hat{c}\hat{d}$	b	$\hat{a} = bcd$	bcd	b	cd
a	$\hat{b}\hat{c}\hat{d}$	a	$\hat{b} = acd$	acd	a	cd
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	0

Ket table showing computation of the commutator $[g,h]$.

For instance, consider the row in the table for the given state: $|acd\rangle = \left|\hat{b}\right\rangle$. The projection operator $P_{a,b} = \{a,b\} \cap ()$ applied to that state yields $\{a\}$ and the projection operator $P_{\hat{a},\hat{b}}$ applied to $|acd\rangle = \left|\hat{b}\right\rangle$ yields $\left\{\hat{a},\hat{b}\right\} \cap \left\{\hat{b}\right\} = \left\{\hat{b}\right\}$ or $\{a,c,d\}$ in the computational basis. Since $P_{a,b}\left(|acd\rangle\right) = |\{a\}\rangle = \left|\left\{\hat{b},\hat{c},\hat{d}\right\}\right\rangle$, the application of $P_{\hat{a},\hat{b}}$ yields $\left\{\hat{b}\right\}$ or $\{a,c,d\}$ in the computational basis. And applying $P_{a,b}$ to $P_{\hat{a},\hat{b}}\left(\hat{b}\right) = \{a,c,d\}$ yields $\{a\}$. Hence the commutator is:

$$[g,h]\left(|\{a,c,d\}\rangle\right) = \left(P_{\hat{a},\hat{b}}P_{a,b} - P_{a,b}P_{\hat{a},\hat{b}}\right)\left(|\{a,c,d\}\rangle\right) = \{a\} - \{a,c,d\} = \{c,d\}$$

or cd in the abbreviated notation of the table.

The indeterminacy principle in full QM connects the commutator to the standard deviations or variances of the probability distributions of the repeated measurements. In QM/sets, each measurement of a projection operator P_S in a state $|T\rangle$ is a Bernoulli trial with the probabilities:

$$\Pr\left(1|T\right) = \frac{|S \cap T|}{|T|} = p \text{ and } \Pr\left(0|T\right) = \frac{|S^c \cap T|}{|T|} = 1 - p.$$

Hence the variance of the probability distribution of measuring P_S in the state T is $var\left(P_S\right)_T = p\left(1-p\right)$. The basic fact behind Heisenberg's indeterminacy or uncertainty principle is that if two operators are incompatible in a given state, then the probability distributions of the two measurements in that state cannot both be sharp. The narrower the distribution for one measurement, the broader the distribution for the measurement of the other operator. In the simple pedagogical model of QM/sets, there is a very simple notion of the probability distribution being sharp, i.e., $var\left(P_S\right)_T = 0$ which means $p = 1$ or $p = 0$. A simple lemma then connects the variances to eigenstates.

Lemma: In a state $T \subseteq U$, if $var\left(P_S\right)_T = 0$, then T is an eigenstate of P_S.

Proof: If $var\left(P_S\right)_T = 0$, then $\Pr\left(1|T\right) = \frac{|S \cap T|}{|T|} = 1$ or 0 so $P_S\left(T\right) = S \cap T = T$ or $P_S\left(T\right) = \emptyset$ so in either case, T is an eigenstate of P_S.□

Corollary: In a state $|T\rangle$, for projection operators P and P', if both $var\left(P\right)_T = 0 = var\left(P'\right)_T$, then $[P,P']\left(|T\rangle\right) = \emptyset$.

Proof: If both variances are 0, then T is a simultaneous eigenstate of the two projections, and thus, by the previous theorem, is in the kernel of the commutator. □

Contrapositing gives what might be taken as the:

Indeterminacy Principle in QM/sets: Given any two projection operators P and P' and a nonzero state $|T\rangle$ where they do not commute, i.e., $[P, P'](|T\rangle) \neq \emptyset$, the two variances cannot both be 0.

In particular, if the given state is an eigenstate of one projection operator (so that variance is 0), then the other (incompatible) projection operator must have a strictly positive variance.

7 Quantum dynamics and the two-slit experiment in QM/sets

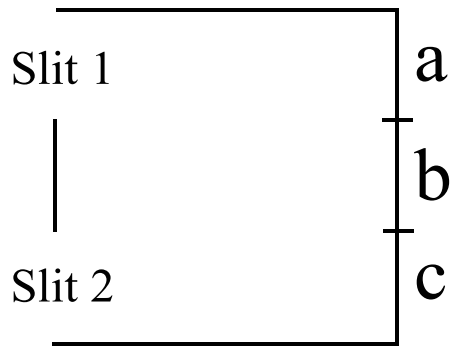
To illustrate a two-slit experiment in quantum mechanics over sets, we need to introduce some "dynamics." In quantum mechanics, the no-distinctions requirement is that the linear transformation has to preserve the degree of indistinctness $\langle\psi|\varphi\rangle$, i.e., that it preserved the inner product. Where two normalized states are fully distinct if $\langle\psi|\varphi\rangle = 0$ and fully indistinct if $\langle\psi|\varphi\rangle = 1$, it is also sufficient to just require that full distinctness and indistinctness be preserved since that would imply orthonormal bases are preserved and that is equivalent to preserving the inner product [13, p. 61]. In QM/sets, we have no inner product but the idea of a linear transformation $A : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ preserving distinctness would simply mean being non-singular. The condition analogous to preserving inner product is $\langle S|_U T\rangle = \langle A(S)|_{A(U)} A(T)\rangle$ where $A(U) = U'$ is defined by $A(\{u\}) = \{u'\}$. For non-singular A, the image $A(U)$ of the U-basis is a basis, i.e., the U'-basis, and the "bracket-preserving" condition holds since $|S \cap T| = |A(S) \cap A(T)|$ for $A(S), A(T) \subseteq A(U) = U'$. Hence the QM/sets analogue of the unitary dynamics of full QM is "non-singular dynamics," i.e., the change-of-state matrix is non-singular.[13]

For $U = \{a, b, c\}$,consider the dynamics: $\{a\} \to \{a, b\}$; $\{b\} \to \{a, b, c\}$; and $\{c\} \to \{b, c\}$ in one time period. This is represented by the non-singular one-period change of state matrix:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

The seven nonzero vectors in the vector space are divided by this "dynamics" into a 4-orbit: $\{a\} \to \{a, b\} \to \{c\} \to \{b, c\} \to \{a\}$, a 2-orbit: $\{b\} \to \{a, b, c\} \to \{b\}$, and a 1-orbit: $\{a, c\} \to \{a, c\}$.

If we take the U-basis vectors as "vertical position" eigenstates, we can device a QM/sets version of the "two-slit experiment" which models "all of the mystery of quantum mechanics" [11, p. 130]. Taking a, b, and c as three vertical positions, we have a vertical diaphragm with slits at a and c. Then there is a screen or wall to the right of the slits so that a "particle" will travel from the diaphragm to the wall in one time period according to the A-dynamics.



[13] In Schumacher and Westmoreland's modal quantum theory [18], they also take the dynamics to be any non-singular linear transformation.

Figure 4: Two-slit setup

We start with or "prepare" the state of a particle being at the slits in the indefinite position state $\{a, c\}$. Then there are two cases.

First case of distinctions at slits: The first case is where we measure the U-state at the slits and then let the resultant position eigenstate evolve by the A-dynamics to hit the wall at the right where the position is measured again. The probability that the particle is at slit 1 or at slit 2 is:

$$\Pr(\{a\} \text{ measured at slits} \mid \{a, c\} \text{ at slits}) = \frac{\langle \{a\} |_U \{a,c\} \rangle^2}{\|\{a,c\}\|_U^2} = \frac{|\{a\} \cap \{a,c\}|}{|\{a,c\}|} = \frac{1}{2};$$
$$\Pr(\{c\} \text{ measured at slits} \mid \{a, c\} \text{ at slits}) = \frac{\langle \{c\} |_U \{a,c\} \rangle^2}{\|\{a,c\}\|_U^2} = \frac{|\{c\} \cap \{a,c\}|}{|\{a,c\}|} = \frac{1}{2}.$$

If the particle was at slit 1, i.e., was in eigenstate $\{a\}$, then it evolves in one time period by the A-dynamics to $\{a, b\}$ where the position measurements yield the probabilities of being at a or at b as:

$$\Pr(\{a\} \text{ measured at wall} \mid \{a, b\} \text{ at wall}) = \frac{\langle \{a\} |_U \{a,b\} \rangle^2}{\|\{a,b\}\|_U^2} = \frac{|\{a\} \cap \{a,b\}|}{|\{a,b\}|} = \frac{1}{2}$$
$$\Pr(\{b\} \text{ measured at wall} \mid \{a, b\} \text{ at wall}) = \frac{\langle \{b\} |_U \{a,b\} \rangle^2}{\|\{a,b\}\|_U^2} = \frac{|\{b\} \cap \{a,b\}|}{|\{a,b\}|} = \frac{1}{2}.$$

If on the other hand the particle was found in the first measurement to be at slit 2, i.e., was in eigenstate $\{c\}$, then it evolved in one time period by the A-dynamics to $\{b, c\}$ where the position measurements yield the probabilities of being at b or at c as:

$$\Pr(\{b\} \text{ measured at wall} \mid \{b, c\} \text{ at wall}) = \frac{|\{b\} \cap \{b,c\}|}{|\{b,c\}|} = \frac{1}{2}$$
$$\Pr(\{c\} \text{ measured at wall} \mid \{b, c\} \text{ at wall}) = \frac{|\{c\} \cap \{b,c\}|}{|\{b,c\}|} = \frac{1}{2}.$$

Hence we can use the laws of probability theory to compute the probabilities of the particle being measured at the three positions on the wall at the right if it starts at the slits in the superposition state $\{a, c\}$ *and* the measurements were made at the slits:

$$\Pr(\{a\} \text{ measured at wall} \mid \{a, c\} \text{ at slits}) = \tfrac{1}{2}\tfrac{1}{2} = \tfrac{1}{4};$$
$$\Pr(\{b\} \text{ measured at wall} \mid \{a, c\} \text{ at slits}) = \tfrac{1}{2}\tfrac{1}{2} + \tfrac{1}{2}\tfrac{1}{2} = \tfrac{1}{2};$$
$$\Pr(\{c\} \text{ measured at wall} \mid \{a, c\} \text{ at slits}) = \tfrac{1}{2}\tfrac{1}{2} = \tfrac{1}{4}.$$

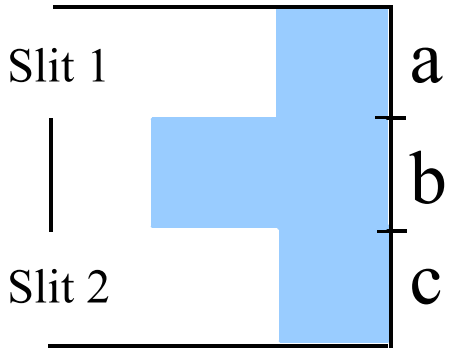


Figure 5: Final probability distribution with measurements at slits

This the QM/sets version of the usual sum of the probability distributions for the particle going through slit 1 or going through slit 2.

Second case of no distinctions at slits: The second case is when no measurements are made at the slits and then the superposition state $\{a, c\}$ evolves by the A-dynamics to $\{a, b\}+\{b, c\} = \{a, c\}$ where the superposition at $\{b\}$ cancels out. Then the final probabilities will just be probabilities of finding $\{a\}$, $\{b\}$, or $\{c\}$ when the measurement is made only at the wall on the right is:

$$\Pr(\{a\} \text{ measured at wall } | \{a, c\} \text{ at slits}) = \Pr(\{a\} | \{a, c\}) = \frac{|\{a\} \cap \{a, c\}|}{|\{a, c\}|} = \frac{1}{2};$$
$$\Pr(\{b\} \text{ measured at wall } | \{a, c\} \text{ at slits}) = \Pr(\{b\} | \{a, c\}) = \frac{|\{b\} \cap \{a, c\}|}{|\{a, c\}|} = 0;$$
$$\Pr(\{c\} \text{ measured at wall } | \{a, c\} \text{ at slits}) = \Pr(\{c\} | \{a, c\}) = \frac{|\{c\} \cap \{a, c\}|}{|\{a, c\}|} = \frac{1}{2}.$$

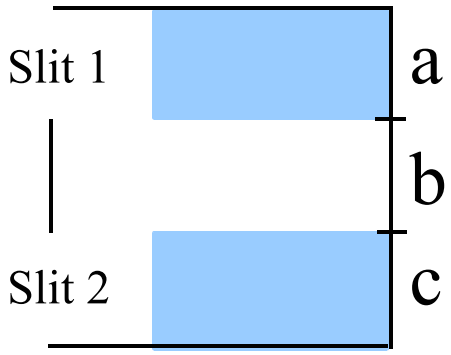


Figure 6: Final probability distribution with no measurement at slits

Since no "collapse" took place at the slits due to no distinctions being made there, the indistinct element $\{a, c\}$ evolved (rather than one or the other of the distinct elements $\{a\}$ or $\{c\}$). The action of A is the same on $\{a\}$ and $\{c\}$ as when they evolve separately since A is a linear operator but the two results are now added together *as part of the evolution*. This allows the "interference" of the two results and thus the cancellation of the $\{b\}$ term in $\{a, b\} + \{b, c\} = \{a, c\}$. The addition is, of course, mod 2 (where $-1 = +1$) so, in "wave language," the "wave crest" and "wave trough" that add at the location $\{b\}$ cancel out. When this indistinct element $\{a, c\}$ "hits the wall" on the right, there is an equal probability of that distinction yielding either of those eigenstates. Figure 6 shows the simplest example of the "light and dark bands" characteristic of superposition and interference illustrating "all of the mystery of quantum mechanics". This pedagogical model gives the simple logical essence of the two-slit experiment without the complex-valued wave functions that distract from the essential point–which is the difference between the separate mixed state evolutions resulting from measurement at the slits, and the combined evolution of the superposition $\{a, c\}$ that allows interference.

8 A Simple Bell Theorem in QM/Sets

A simple version of a Bell inequality can be derived in the case of $\mathbb{Z}_2^2$ with three bases $A = \{A^+, A^-\}$, $B = \{B^+, B^-\}$, and $C = \{C^+, C^-\}$ (like a particle having spin up or down along three different axes), and where the kets are:

kets	A-basis	B-basis	C-basis
$\|1\rangle$	$\{A^+, A^-\}$	$\{B^+\}$	$\{C^+\}$
$\|2\rangle$	$\{A^-\}$	$\{B^-\}$	$\{C^+, C^-\}$
$\|3\rangle$	$\{A^+\}$	$\{B^+, B^-\}$	$\{C^-\}$
$\|4\rangle$	$\emptyset$	$\emptyset$	$\emptyset$

Ket table for $\wp(A) \cong \wp(B) \cong \wp(C) \cong \mathbb{Z}_2^2$.

Given a ket in $\mathbb{Z}_2^2 \cong \wp(A) \cong \wp(B) \cong \wp(C)$, and using the usual equiprobability assumption on sets, the probabilities of getting the different outcomes for the various "observables" in the different given states are given in the following table.

Given state \ Outcome of test	A^+	A^-	B^+	B^-	C^+	C^-
$\{A^+, A^-\} = \{B^+\} = \{C^+\}$	$\frac{1}{2}$	$\frac{1}{2}$	1	0	1	0
$\{A^-\} = \{B^-\} = \{C^+, C^-\}$	0	1	0	1	$\frac{1}{2}$	$\frac{1}{2}$
$\{A^+\} = \{B^+, B^-\} = \{C^-\}$	1	0	$\frac{1}{2}$	$\frac{1}{2}$	0	1

State-outcome table.

The tensor product of two state spaces, e.g., $\wp(A) \otimes \wp(A)$, is the space generated by the ordered pairs $A \times A$, i.e., $\wp(A \times A)$, Thus in the A-basis, the basis elements are the elements of $A \times A$ and the vectors are all the subsets in $\wp(A \times A)$. But we could obtain the same space as $\wp(B \times B)$ and $\wp(C \times C)$, and we can construct a ket table where each row is a ket expressed in the different bases. And these calculations in terms of sets could also be carried out in terms of vector spaces over $\mathbb{Z}_2$ where the rows of the ket table are the kets in the tensor product:

$$\mathbb{Z}_2^2 \otimes \mathbb{Z}_2^2 \cong \wp(A \times A) \cong \wp(B \times B) \cong \wp(C \times C).$$

Since $\{A^+\} = \{B^+, B^-\} = \{C^-\}$ and $\{A^-\} = \{B^-\} = \{C^+, C^-\}$, the subset $\{A^+\} \times \{A^-\} = \{(A^+, A^-)\} \subseteq A \times A$ is expressed in the $B \times B$-basis as $\{B^+, B^-\} \times \{B^-\} = \{(B^+, B^-), (B^-, B^-)\}$, and in the $C \times C$-basis it is $\{C^-\} \times \{C^+, C^-\} = \{(C^-, C^+), (C^-, C^-)\}$. Hence one row in the ket table (the second row) has:

$$\{(A^+, A^-)\} = \{(B^+, B^-), (B^-, B^-)\} = \{(C^-, C^+), (C^-, C^-)\}.$$

Since the full ket table has 16 rows, we will just give a partial table that suffices for our calculations.

$A \times A$	$B \times B$	$C \times C$
$\{(A^+, A^-)\}$	$\{(B^+, B^-), (B^-, B^-)\}$	$\{(C^-, C^+), (C^-, C^-)\}$
$\{(A^-, A^+)\}$	$\{(B^-, B^+), (B^-, B^-)\}$	$\{(C^+, C^-), (C^-, C^-)\}$
$\{(A^+, A^+), (A^+, A^-)\}$	$\{(B^+, B^+), (B^-, B^+)\}$	$\{(C^-, C^+)\}$
$\{(A^+, A^+), (A^-, A^-)\}$	$\{(B^+, B^+), (B^+, B^-), (B^-, B^+)\}$	$\{(C^+, C^+), (C^+, C^-), (C^-, C^+)\}$
$\{(A^+, A^-), (A^-, A^+)\}$	$\{(B^+, B^-), (B^-, B^+)\}$	$\{(C^+, C^-), (C^-, C^+)\}$

Partial ket table for $\wp(A \times A) \cong \wp(B \times B) \cong \wp(C \times C)$

We can classify each vector or subset as "separated" or "entangled" and we can furthermore see how that is independent of the basis. For instance $\{(A^+, A^+), (A^+, A^-)\}$ is "separated" since:

$$\begin{aligned}\{(A^+, A^+), (A^+, A^-)\} &= \{A^+\} \times \{A^+, A^-\} = \{(B^+, B^+), (B^-, B^+)\} \\ &= \{B^+, B^-\} \times \{B^+\} = \{(C^-, C^+)\} = \{C^-\} \times \{C^+\}.\end{aligned}$$

An example of an "entangled state" is:

$$\begin{aligned}&\{(A^+, A^-), (A^-, A^+)\} \\ &= \{(B^+, B^-), (B^-, B^-), (B^-, B^+), (B^-, B^-)\} = \{(B^+, B^-), (B^-, B^+)\} \\ &= \{(C^-, C^+), (C^-, C^-), (C^+, C^-), (C^-, C^-)\} = \{(C^+, C^-), (C^-, C^+)\}.\end{aligned}$$

Taking this entangled "Bell state" as the initial state (last row in the ket table), there is a probability distribution on $A \times B \times C$ where $\Pr(A^+, B^+, C^+)$ (for instance) is defined as the probability of getting the result $\{A^+\}$ if a A-measurement is performed on the left-hand system, and if instead a B-measurement is performed on the left-hand system then $\{B^+\}$ is obtained, and if instead a C-measurement is performed on the left-hand system then $\{C^+\}$ is obtained. Thus we would have $\Pr(A^+, B^+, C^+) = \frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{8}$. In this way the probability distribution $\Pr(x, y, z)$ is defined on $A \times B \times C$.

A Bell inequality can be obtained from this joint probability distribution over the outcomes $A \times B \times C$ of measuring these three incompatible attributes [5]. Consider the following marginals:

$$\Pr\left(A^+, B^+\right) = \Pr\left(A^+, B^+, C^+\right) + \Pr\left(A^+, B^+, C^-\right) \checkmark$$
$$\Pr\left(B^-, C^-\right) = \Pr\left(A^+, B^-, C^-\right) \checkmark + \Pr\left(A^-, B^-, C^-\right)$$
$$\Pr\left(A^+, C^-\right) = \Pr\left(A^+, B^+, C^-\right) \checkmark + \Pr\left(A^+, B^-, C^-\right) \checkmark.$$

The two terms in the last marginal are each contained in one of the two previous marginals (as indicated by the check marks) and all the probabilities are non-negative, so we have the following inequality:

$$\Pr(A^+, B^+) + \Pr(B^-, C^-) \geq \Pr(A^+, C^-)$$
Bell inequality.

All this has to do with measurements on the left-hand system. But there is an alternative interpretation to the probabilities $\Pr(x, y)$, $\Pr(y, z)$, and $\Pr(x, z)$ *if* we assume that the outcome of a measurement on the right-hand system is *independent* of the outcome of the same measurement on the left-hand system. Then $\Pr(A^+, B^+)$ is the probability of a A-measurement on the left-hand system giving $\{A^+\}$ and then a B-measurement on the right-hand system giving $\{B^+\}$, and so forth. Under that *independence assumption* and for this initially prepared "Bell state" (which is left-right symmetrical in each basis),

$$\{(A^+, A^-), (A^-, A^+)\} = \{(B^+, B^-), (B^-, B^+)\} = \{(C^+, C^-), (C^-, C^+)\},$$

the probabilities would be the same. That is, under that assumption, the probabilities, $\Pr(A^+) = \frac{1}{2} = \Pr(A^-)$, $\Pr(B^+) = \frac{1}{2} = \Pr(C^+)$, and $\Pr(B^-) = \frac{1}{2} = \Pr(C^-)$ are the same regardless of whether we are measuring the left-hand or right-hand system of that composite state. Hence the above Bell inequality would still hold. But we can use QM/Sets to compute the probabilities for those different measurements on the two systems to see if the independence assumption is compatible with QM/Sets.

To compute $\Pr(A^+, B^+)$, we first measure the left-hand component in the A-basis. Since the given state is $\{(A^+, A^-), (A^-, A^+)\}$, and (A^+, A^-) and (A^-, A^+) are equiprobable, the probability of getting $\{A^+\}$ (i.e., the eigenvalue 1 for the observable $\chi_{\{A^+\}}$) is $\frac{1}{2}$. But the right-hand system is then in the state $\{A^-\}$ and the probability of getting $\{B^+\}$ (i.e., eigenvalue 0 for the observable $\chi_{\{B^+\}}$) is 0 (as seen in the state-outcome table). Thus the probability is $\Pr(A^+, B^+) = \frac{1}{2}0 = 0$.

To compute $\Pr(B^-, C^-)$, we first perform a B-basis "measurement" on the left-hand component of the given state $\{(A^+, A^-), (A^-, A^+)\} = \{(B^+, B^-), (B^-, B^+)\}$, and we see that the probability of getting $\{B^-\}$ is $\frac{1}{2}$. Then the right-hand system is in the state $\{B^+\}$ and the probability of getting $\{C^-\}$ in a C-basis "measurement" of the right-hand system in the state $\{B^+\}$ is 0 (as seen from the state-outcome table). Hence the probability is $\Pr(B^-, C^-) = 0$.

Finally we compute $\Pr(A^+, C^-)$ by first making an A-measurement on the left-hand component of the given state $\{(A^+, A^-), (A^-, A^+)\}$ and get the result $\{A^+\}$ with probability $\frac{1}{2}$. Then the state of the second system is $\{A^-\}$ so a C-measurement will give the $\{C^-\}$ result with probability $\frac{1}{2}$ so the probability is $\Pr(A^+, C^-) = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$.

Then we plug the probabilities into the Bell inequality:

$$\Pr(A^+, B^+) + \Pr(B^-, C^-) \geq \Pr(A^+, C^-)$$
$$0 + 0 \ngeq \tfrac{1}{4}$$
Violation of Bell inequality.

The violation of the Bell inequality shows that the independence assumption about the measurement outcomes on the left-hand and right-hand systems is incompatible with QM/Sets so the effects of the measurements are "nonlocal."

9 Identical particles in QM/Sets

In QM/Sets, the tensor product simplifies to the Cartesian product in the sense that: $\wp(U) \otimes \wp(U') \cong \wp(U \times U')$ via the mapping $\{u\} \otimes \{u'\} \longleftrightarrow \{(u, u')\}$ between basis sets. Suppose we have two identical (indistinguishable) particles which we (following Weyl) artificially label "Mike" and "Ike." Each particle can be in one of three eigenstates A, B, and C so the single particle state space $\wp(U)$ is generated by $U = \{A, B, C\}$ and thus the space of two-particle states is generated by the nine states in $U \times U = \{A, B, C\} \times \{A, B, C\}$.

We define an equivalence relation on the nine basis elements where each ordered pair is equivalent to the one where Mike and Ike are permuted. There are the following six equivalence classes on the basis elements. If we take the basis set $U \times U$ as the sample space with each pair as being equiprobable, then the probability of the equivalence classes is the *Maxwell-Boltzmann distribution.*

Equivalence classes under permutation	M-B
$\{(A, B), (B, A)\}$	$\frac{2}{9}$
$\{(A, C), (C, A)\}$	$\frac{2}{9}$
$\{(B, C), (C, B)\}$	$\frac{2}{9}$
$\{(A, A)\}$	$\frac{1}{9}$
$\{(B, B)\}$	$\frac{1}{9}$
$\{(C, C)\}$	$\frac{1}{9}$

Maxwell-Boltzmann distribution

In quantum mechanics, it is often said that the result obtained after permuting the particles is "indistinguishable" from the prior state, but that is also true for classical particles. It would be more accurate to say that the result in the quantum case is *identical* when the artificially labelled particles are permuted. In other words, what is classically seen as an equivalence class of numerically-distinct states obtained by permutations should really be seen in the quantum case as one "indefinite" entity that is identical under permutation. In that case, the sample space has six indefinite entities which we may take as equiprobable to obtain the *Bose-Einstein distribution* and they span a six-dimensional subspace V_{BE} of the nine-dimensional $\wp(U \times U)$.

Six indefinite states	B-E
$\{(A, B), (B, A)\}$	$\frac{1}{6}$
$\{(A, C), (C, A)\}$	$\frac{1}{6}$
$\{(B, C), (C, B)\}$	$\frac{1}{6}$
$\{(A, A)\}$	$\frac{1}{6}$
$\{(B, B)\}$	$\frac{1}{6}$
$\{(C, C)\}$	$\frac{1}{6}$

Bose-Einstein distribution

In enumerative combinatorial theory, there is a basic distinction between allowing repetitions and not allowing repetitions. We have been implicitly not allowing repetitions so that applying the Mike-Ike permutation to (A, A) since we only listed it once to obtain the above six B-E states.

But allowing repetitions, we get another (A, A) in the superposition state $\{(A, A), (A, A)\} = 0$. Thus allowing repetitions gives only three non-zero indefinite states which form the sample space of equiprobable points for the Fermi-Dirac distribution and which span a three-dimensional subspace V_{FD} of the nine-dimensional $\wp(U \times U)$.

Three indefinite states	F-D
$\{(A, B), (B, A)\}$	$\frac{1}{3}$
$\{(A, C), (C, A)\}$	$\frac{1}{3}$
$\{(B, C), (C, B)\}$	$\frac{1}{3}$
$\{(A, A), (A, A)\} = 0$	0
$\{(B, B), (B, B)\} = 0$	0
$\{(C, C), (C, C)\} = 0$	0

Fermi-Dirac distribution

Since these computations in QM/Sets are essentially combinatorial, the boson-fermion distinction illustrates on "the combinatorial level, the duality between... balls into boxes (subject to certain conditions) not allowing repetitions or allowing repetitions." [20, p. 295] When repetitions are allowed (so they cancel out), then the Pauli exclusion principle is a consequence of addition mod 2 whereas in ordinary QM, it is a consequence of treating transpositions as being antisymmetric.

10 Concluding remarks

This paper is intended only as an introduction to the pedagogical model of QM/sets. Enough of the pedagogical model has been developed to show how it hangs together and how it may elucidate some of the more perplexing aspects of full QM by seeing them in a simple setting.

11 Appendix: Transporting Vector Space Structures

It is important to rigorously understand the mathematics connecting finite-dimensional QM over $\mathbb{C}^n$ to QM/Sets over $\mathbb{Z}_2^n$. There is a general method to transport some structures from a vector space V over a field $\mathbb{K}$ to a vector space V' over a *different* field $\mathbb{K}'$. Select a basis set U for the source space V and then consider a structure on V that can be characterized in terms of the basis set U. Then apply the free vector space over the field $\mathbb{K}'$ construction to U to generate the target vector space V'. Since the source structure was defined in terms of the basis set U, it can be carried over or "transported" to V' via *its* basis set U.

This method can be stated in rigorous terms using category theory ([15]; [2]). The construction of the free vector space over a field $\mathbb{K}$ is a functor from the category $Sets$ of sets and functions to the category $Vect_{\mathbb{K}}$ of vector spaces over $\mathbb{K}$ and linear transformations. The functor will only be used here on finite sets where it takes a finite set U to the vector space $\mathbb{K}^U$. The primary structures being transported are direct-sum decompositions (DSD) of a finite-dimensional vector space V. A DSD a set $\{V_i\}$ of disjoint subspaces (i.e., only overlap is zero space) so that the whole space V is their direct sum, or, in terms of category theory, V is the coproduct $V = \oplus V_i$ of the subspaces $\{V_i\}$. In the category $Sets$, a set $\{B_i\}$ of disjoint subsets of a set U is a set partition of U if $\cup B_i = U$, or, in terms of category theory, U is the coproduct of the disjoint subsets $\{B_i\}$. The free vector space over $\mathbb{K}$ functor is a left adjoint, "left adjoints preserve colimits" [2, p. 197], and coproducts are a special type of colimit. Hence the free vector space functor carries a set partition $\pi = \{B_i\}_{i=1,...,m}$ to the DSD $\left\{V_i = \mathbb{K}^{B_i}\right\}$ of $V = \mathbb{K}^U = \oplus \mathbb{K}^{B_i}$.

Now start with the structure of a DSD $\{V_i\}$ on $V \in Vect_{\mathbb{K}}$. What we previously called "characterizing the structure in terms of a basis set U" is rigorously interpreted to mean, in this case, finding a basis U and a partition $\{B_i\}$ on U so that the given DSD $\{V_i\}$ is the image of the free

vector space functor, i.e., $V = \mathbb{K}^U = \oplus \mathbb{K}^{B_i} = \oplus V_i$. But then the free vector space functor over a different field $\mathbb{K}'$ can be applied to the same set partition $\{B_i\}$ of the set U to generate a DSD $\left\{V_i' = \mathbb{K}'^{B_i}\right\}$ of $V' = \mathbb{K}'^U$. That is how to rigorously describe "transporting" a set-based structure on a vector V over $\mathbb{K}$ to a vector space V' over a different field $\mathbb{K}'$.

To show that any given DSD $\{V_i\}$ of V is in the image of the free vector space over $\mathbb{K}$ functor, pick basis set B_i of V_i. The sets B_i are disjoint and since $\{V_i\}$ is a DSD, the union $U = \cup B_i$ is a basis for V so $V_i = \mathbb{K}^{B_i}$ and $V = \mathbb{K}^U = \oplus \mathbb{K}^{B_i}$.

This method is applied to the transporting of self-adjoint operators from $V = \mathbb{C}^n$ to $V' = \mathbb{Z}_2^n$ that motivates QM/Sets. A self-adjoint operator $F : \mathbb{C}^n \to \mathbb{C}^n$ has a basis $U = \{u_1, ..., u_n\}$ of orthonormal eigenvectors and it has real distinct eigenvalues $\{\phi_i\}_{j=1,...,m}$, so it defines the real eigenvalue function $f : U \to \mathbb{R}$ where for $u_j \in U$, $f(u_j)$ is one of the distinct eigenvalues $\{\phi_i\}_{i=1,...m}$. For each distinct eigenvalue ϕ_i, there is the eigenspace V_i of its eigenvectors and $\{V_i\}_{i=1,...,m}$ is a DSD on $V = \mathbb{C}^n$. The inverse-image $\pi = \left\{B_i = f^{-1}(\phi_i)\right\}_{i=1,...,m}$ of the eigenvalue function $f : U \to \mathbb{R}$ is a set partition on U.

Thus the set-based structure we have is the set U with a partition $\left\{B_i = f^{-1}(\phi_i)\right\}_i$ on U induced by a real-value function $f : U \to \mathbb{R}$ on U. That set-based structure is sufficient to reconstruct the DSD $\left\{V_i = \mathbb{C}^{B_i}\right\}_i$ on $V = \mathbb{C}^n \cong \mathbb{C}^U = \oplus \mathbb{C}^{B_i}$ as well as the original operator F. The operator F is defined on the basis U by $Fu_j = f(u_j)\, u_j$ for $j = 1, ..., n$. That process of going from the function $f : U \to \mathbb{R}$ on a basis set U of $\mathbb{C}^U$ to an operator on $\mathbb{C}^U$ might be called *internalizing* the function $f : U \to \mathbb{R}$ in $\mathbb{C}^U$.

Given the set-based structure of a real-valued function $f : U \to \mathbb{R}$, which determines the set partition $\left\{f^{-1}(\phi_i)\right\}_{i=1,...,m}$ on U, we then apply the free vector space over $\mathbb{Z}_2$ functor to construct the vector space $\mathbb{Z}_2^U$. That vector space is more familiar in the form of the powerset $\wp(U) \cong \mathbb{Z}_2^U$ since each function $U \to \mathbb{Z}_2 = \{0, 1\}$ in $\mathbb{Z}_2^U$ is the characteristic function χ_S of a subset $S \in \wp(U)$. The free vector space functor $\mathbb{Z}_2^{()}$ takes the coproduct $U = \cup_{i=1}^m f^{-1}(\phi_i)$ to the DSD $\left\{\wp\left(f^{-1}(\phi_i)\right)\right\}$ of $\wp(U)$. The attempt to internalize the real function $f : U \to \mathbb{R}$ would only work if f took values in $\mathbb{Z}_2 = \{0, 1\} \subseteq \mathbb{R}$ in which case f would be a characteristic function χ_S for some subset $S \in \wp(U)$. In that special case, the internalized operator would be the projection operator $P_S : \mathbb{Z}_2^U \to \mathbb{Z}_2^U$ which in terms of the basis U has the action $P_S(T) = S \cap T$ taking any subset $T \in \wp(U)$ to $S \cap T \in \wp(S)$.

Hence outside of characteristic functions, the real-valued functions $f : U \to \mathbb{R}$ cannot be internalized as operators on $\mathbb{Z}_2^U$. But that is fine since the idea of the model QM/Sets is that given a basis U of $\mathbb{Z}_2^n$, the quantum probability calculus will just be the classical finite probability calculus with the outcome set or sample space U where $f : U \to \mathbb{R}$ is a real-valued random variable. We have illustrated the transporting of set-based structures on $\mathbb{C}^n$ to $\mathbb{Z}_2^n$ using a basis set U, but in the stand-alone model QM/Sets, we cut the umbilical cord to $\mathbb{C}^n$ and work with any other basis U' of $\mathbb{Z}_2^n$ and real-valued random variables $g : U' \to \mathbb{R}$ on *that* sample space.

Other structures can be transported across the bridge from $\mathbb{C}^n$ to $\mathbb{Z}_2^n$. QM/Sets differs from the other four attempts to define some toy version of QM on sets by the treatment of the Dirac brackets. Starting with our orthonormal basis U on a finite-dimensional Hilbert space $\mathbb{C}^n$ (where the bracket is the inner product), we need to define the transported brackets applied to two subsets $S, T \subseteq U$ in $\wp(U)$. The two subsets define the vectors $\psi_S = \sum_{u \in S} |u\rangle$ and $\psi_T = \sum_{u \in T} |u\rangle$ in $\mathbb{C}^n$ which have the bracket value $\langle \psi_S | \psi_T \rangle = |S \cap T|$. Since that value is defined just in terms of the subsets $S, T \subseteq U$ as the cardinality of their overlap, that value can be transported to $\wp(U)$ as the real-valued basis-dependent brackets $\langle S|_U T \rangle = |S \cap T|$.

References

[1] Abramsky, Samson and Coecke, Bob 2004. A categorical semantics of quantum protocols. in: *Proceedings of the 19th IEEE Symposium on Logic in Computer Science (LiCS'04)*, IEEE Computer Science Press, 415–425.

[2] Awodey, Steve. 2006. *Category Theory*. Oxford: Clarendon Press.

[3] Boole, George 1854. *An Investigation of the Laws of Thought on which are founded the Mathematical Theories of Logic and Probabilities*. Cambridge: Macmillan and Co.

[4] Cohen-Tannoudji, Claude, Bernard Diu and Franck Laloë 2005. *Quantum Mechanics Vol. 1*. New York: John Wiley & Sons.

[5] D'Espagnat, Bernard. 1979. The Quantum Theory and Reality. *Scientific American* 241 (5): 158–81.

[6] Dirac, Paul A. M. 1958. *The Principles of Quantum Mechanics (4th ed.)*. Oxford: Clarendon.

[7] Eddington, Arthur S. 1947. *New Pathways in Science (Messenger Lectures 1934)*. Cambridge UK: Cambridge University Press.

[8] Ellerman, David 2009. Counting Distinctions: On the Conceptual Foundations of Shannon's Information Theory. *Synthese*. 168 (1 May): 119-149.

[9] Ellerman, David 2010. The Logic of Partitions: Introduction to the Dual of the Logic of Subsets. *Review of Symbolic Logic*. 3 (2 June): 287-350.

[10] Ellerman, David 2014. An Introduction of Partition Logic. *Logic Journal of the IGPL*. 22, no. 1: 94–125.

[11] Feynman, Richard P. 1967. *The Character of Physical Law*. Cambridge: MIT Press.

[12] Hanson, Andrew J., Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai 2013. *Discrete Quantum Theories*. arXiv:1305.3292v1.

[13] Hoffman, K., and R. Kunze. 1961. *Linear Algebra*. Englewood Cliffs NJ: Prentice-Hall.

[14] Laplace, Pierre-Simon. 1995 (1825). *Philosophical Essay on Probabilities*. Translated by A. I. Dale. New York: Springer-Verlag.

[15] Mac Lane, Saunders. 1998. *Categories for the Working Mathematician. 2nd Ed*. New York: Springer Science+Business Media.

[16] McEliece, Robert J. 1977. *The Theory of Information and Coding: A Mathematical Framework for Communication (Encyclopedia of Mathematics and its Applications, Vol. 3)*. Reading MA: Addison-Wesley.

[17] Nielsen, M., and I. Chuang. 2000. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.

[18] Schumacher, Benjamin and Michael Westmoreland 2012. Modal Quantum Theory. *Foundations of Physics*, 42, 918-925.

[19] Spekkens, Robert W. 2007. Evidence for the Epistemic View of Quantum States: A Toy Theory. *Physical Review A* 75 (3): 032110.

[20] Stanley, Richard P. 1999. *Enumerative Combinatorics Vol. 2*. New York: Cambridge University Press.

[21] Takeuchi, T., L.N. Chang, Z. Lewis, and D. Minic. 2012. *Some Mutant Forms of Quantum Mechanics*. [quant-Ph] arXiv:1208.5544v1.

[22] Weyl, Hermann 1949. *Philosophy of Mathematics and Natural Science*. Princeton NJ: Princeton University Press.