Quantum mechanics over sets

David Ellerman
Department of Philosophy
U. of California/Riverside

November 9, 2013

Abstract

In the tradition of toy models of quantum mechanics in vector spaces over finite fields (e.g., Schumacher and Westmoreland's "modal quantum theory"), one finite field stands out, $\mathbb{Z}_2$, since vectors over $\mathbb{Z}_2$ have an interpretation as natural mathematical objects, i.e., sets. This engages a sets-to-vector-spaces bridge that is part of the mathematical folklore to translate both ways between set concepts and vector space concepts. Using that bridge, the mathematical framework of (finite-dimensional) quantum mechanics can be transported down to sets resulting in *quantum mechanics over sets* or *QM/sets*. This approach leads to a different treatment of Dirac's brackets than in "modal quantum theory" (MQT), and that gives a full probability calculus (unlike MQT that only has zero-one modalities of impossible and possible). That, in turn, leads to a rather fulsome theory of QM over sets that includes "logical" models of the double-slit experiment, Bell's Theorem, quantum information theory, quantum computing, and much else. Indeed, QM/sets is proposed as the "logic" of QM in the old-fashioned sense of "logic" as giving the simplified essentials of a theory. QM/sets is also a key part of a broader research program to provide an interpretation of QM based on the notion of "objective indefiniteness," a program that grew out the recent development of the logic of partitions mathematically dual to the usual Boolean logic of subsets.

Contents

Part I
The lifting program and the probability calculus

1 Toy models of QM over finite fields

In the tradition of "toy models" for quantum mechanics (QM), Schumacher and Westmoreland [21], Hanson et al. [14], and Takeuchi, Chang, et al. [24] [5], have recently investigated models of quantum mechanics over finite fields. One finite field stands out over the rest, $\mathbb{Z}_2$, since vectors in a vector space over $\mathbb{Z}_2$ have a natural interpretation, namely as *sets* that are subsets of a universe set. But

in any vector space over a finite field, there is no inner product so the first problem in constructing a toy model of QM in this context is the definition of Dirac's brackets. Which aspects of the usual treatment of the brackets should be retained and which aspects should be dropped?

Schumacher and Westmoreland (S&W) chose to have their brackets continue to have values in the base field, e.g., $\mathbb{Z}_2 = \{0, 1\}$, so their "theory does not make use of the idea of probability."[21, p. 919] Instead, the values of 0 and 1 are respectively interpreted modally as *impossible* and *possible* and hence their name of "modal quantum theory." A number of results from full QM carry over to their modal quantum theory, e.g., no-cloning, superdense coding, and teleportation, but without a probability calculus, other results such as Bell's Theorem do not carry over: "in the absence of probabilities and expectation values the Bell approach will not work." [21, p. 921] Hence they develop a variation using the modal concepts from a toy model by Hardy. [15][1]

But all these limitations can be overcome by the different treatment of the brackets taken here which yields a full probability calculus for a model of *quantum mechanics over sets* (QM/sets) using the $\mathbb{Z}_2$ base field. Binary coding theory also uses vector spaces over $\mathbb{Z}_2$, and one of the principal functions, the Hamming distance function [19], takes non-negative integer values. Applied to two subsets S, T of a given universe set U, the Hamming distance function is the cardinality $|S + T|$ of their symmetric difference (i.e., the number of places where the two binary strings differ). In full QM, the bracket $\langle\psi|\varphi\rangle$ is taken as the size of the "overlap" between the two states. Hence it is natural in QM/sets to define the bracket $\langle S|T\rangle$ applied to subsets $S, T \subseteq U$ as the size of their overlap, i.e., the cardinality $|S \cap T|$ of their intersection.

The usual QM formalism (always finite dimensional), e.g., the norm as the square root of the brackets $|\psi| = \sqrt{\langle\psi|\psi\rangle}$, can be developed in this context, and then Born's Rule yields a probability calculus. And it is essentially a familiar calculus, logical probability theory for a finite universe set of outcomes developed by Laplace, Boole, and others. The only difference from that classical calculus is the vector space formulation which allows different (equicardinal) bases or universe sets of outcomes and thus it is "non-commutative." This allows the development of the QM/sets version of many QM results such as Bell's Theorem, the indeterminacy principle, double-slit experiments, and much else in the context of finite sets. And that, in turn, helps to illuminate some of the seemingly "weird" aspects of full QM.

By developing a sets-version of QM, the concepts and relationships of full QM are represented in a pared-down ultra-simple version that can be seen as representing the essential "logic" of QM. It represents the "logic of QM" in that old sense of "logic" as giving the basic essentials of a theory (even reduced to "zero-oneness"), not in the sense of giving the behavior of propositions in a theory (which is the usual "quantum logic"). This approach to full QM [11] arises out of the recent development of the logic of partitions ([10] and [12]) that is (category-theoretically) dual to the ordinary Boolean logic of subsets (which is usually mis-specified as the special case of propositional logic).

2 The lifting sets-to-vector-spaces program

2.1 The basis principle

There is a natural bridge (or ladder) between QM/sets and full QM based on the mathematical relation between sets and vector spaces that is part of the mathematical folklore. A subset can be viewed as a vector in a vector space over $\mathbb{Z}_2$, and a vector expressed in a basis can be viewed as a linearized set where each (basis-) element in the set has a coefficient in the base field of scalars. Using this conceptual bridge (or ladder), set-based concepts as in QM/sets can be transported or "lifted" to vector space concepts as in QM, and vector space concepts may be "delifted" or transported back to set concepts. QM/sets is the delifted version of the mathematical machinery of QM, and, conversely,

[1] Similar remarks apply to the other aforementioned toy models all of which have the brackets taking values in the base field.

the machinery of QM/sets lifts to give the mathematics of QM (but, of course, not the specifically physical assumptions such as the Hamiltonian or the DeBroglie relations connecting energy and frequency or momentum and wavelength).

The bridge from set concepts to vector space concepts has the guiding:

Basis Principle:

Apply the set concept to a basis set and then linearly generate the lifted vector space concept.[2]

For instance, what is the vector space lift of the set concept of cardinality? We apply the set concept of cardinality to a basis set of a vector space where it yields the notion of *dimension* of the vector space (after checking that all bases have equal cardinality). Thus the lift of set-cardinality is not the cardinality of a vector space but its dimension.[3] Thus the null set $\emptyset$ with cardinality 0 lifts to the trivial zero vector space with dimension 0.

2.2 Lifting partitions to vector spaces

Given a universe set U, a *partition* π of U is a set of non-empty subsets or blocks (or cells) $\{B\}$ of U that are pairwise disjoint and whose union is U. In category-theoretic terms, a partition is a direct sum decomposition of a set, and that concept will lift, in the sets-to-vector-spaces lifting program, to the concept of a direct sum decomposition of a vector space. We obtain this lifting by applying the basis principle. Apply a set partition to a basis set of a vector space. Each block B of the set partition of the basis set linearly generates a subspace $W_B \subseteq V$, and the subspaces together form a *direct sum decomposition*: $V = \sum_B \oplus W_B$. Thus the proper lifted notion of a partition for a vector space is *not* a set partition of a space compatible with the vector space structure as would be defined by a subspace $W \subseteq V$ where $v \sim v'$ if $v - v' \in W$. A *vector space partition* is a direct sum decomposition of the vector space–which is not at all a set partition of the vector space.

2.3 Lifting partition joins to vector spaces

The main partition operation from partition logic that we need to lift to vector spaces is the join operation. Two set partitions cannot be joined unless they are *compatible* in the sense of being defined on the same universe set. This notion of compatibility lifts to vector spaces, via the basis principle, by defining two vector space partitions (i.e., two direct sum decompositions) $\omega = \{W_\lambda\}$ and $\xi = \{X_\mu\}$ on V as being *compatible* if there is a basis set for V so that the two vector space partitions arise from two set partitions of that common or simultaneous basis set.

If two set partitions $\pi = \{B\}$ and $\sigma = \{C\}$ are compatible, then their *join* $\pi \vee \sigma$ is defined as the set partition whose blocks are the non-empty intersections $B \cap C$. Similarly the lifted concept is that if two vector space partitions $\omega = \{W_\lambda\}$ and $\xi = \{X_\mu\}$ are compatible, then their *join* $\omega \vee \xi$ is defined as the vector space partition whose subspaces are the non-zero intersections $W_\lambda \cap X_\mu$. And by the definition of compatibility, we could generate the subspaces of the join $\omega \vee \xi$ by the blocks in the join of the two set partitions of the common basis set.

2.4 Lifting numerical attributes to linear operators

A set partition might be seen as an abstract rendition of the inverse image partition $\left\{f^{-1}(r)\right\}$ defined by some concrete numerical attribute $f : U \to \mathbb{R}$ on U. What is the lift of an attribute? At

[2] Intuitions can be guided by the linearization map which takes a set U to the (free) vector space $\mathbb{C}^U$ where $u \in U$ lifts to the basis vector $\delta_u = \chi_{\{u\}} : U \to \mathbb{C}$. But some choices are involved in the lifting program. For instance, the set attribute $f : U \to \mathbb{R}$ could be taken as defining the linear functional $\mathbb{C}^U \to \mathbb{C}$ that takes δ_u to $f(u)$ or the linear operator $\mathbb{C}^U \to \mathbb{C}^U$ that takes δ_u to $f(u)\,\delta_u$. We will see that the latter is the right choice.

[3] In QM, the extension of concepts on finite dimensional Hilbert space to infinite dimensional ones is well-known. Since our expository purpose is conceptual rather than mathematical, we will stick to finite dimensional spaces.

first glance, the basis principle would seem to imply: define a set numerical attribute on a basis set (with values in the base field) and then linearly generate a functional from the vector space to the base field. But a functional does not define a vector space partition; it only defines the set partition of the vector space compatible with the vector space operations that is determined by the kernel of the functional. Hence we need to try a more careful application of the basis principle.

It is helpful to first give a suggestive reformulation of a set attribute $f : U \to \mathbb{R}$. If f is constant on a subset $S \subseteq U$ with a value r, then we might symbolize this as:

$$f \upharpoonright S = rS$$

and suggestively call S an "eigenvector" and r an "eigenvalue." The multiplication rS is only formal and should be read as: the function f has the value r on the subset S. For any "eigenvalue" r, define power set $\wp(f^{-1}(r))$ = "eigenspace of r" as the set of all the "eigenvectors" with that "eigenvalue." Since the "eigenspaces" span the set U, the attribute $f : U \to \mathbb{R}$ can be represented by:

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

[where $\chi_{f^{-1}(r)}$ is the characteristic function for the set $f^{-1}(r)$ and where the index r runs over the image or "spectrum" of the function $f : U \to \mathbb{R}$].[4] Thus a set attribute determines a set partition and has a constant value on the blocks of the set partition, so by the basis principle, that lifts to a vector space concept that determines a vector space partition and has a constant value on the blocks of the vector space partition.

The suggestive terminology gives the lift. The lift of $f \upharpoonright S = rS$ is the eigenvector equation $Lv = \lambda v$ where L is a linear operator on V. The lift of r is the eigenvalue λ and the lift of an S such that $f \upharpoonright S = rS$ is an eigenvector v such that $Lv = \lambda v$. The lift of an "eigenspace" $\wp(f^{-1}(r))$ is the eigenspace W_λ of an eigenvalue λ. The lift of the simplest attributes, which are the characteristic functions $\chi_{f^{-1}(r)}$, are the projection operators P_λ that project to the eigenspaces W_λ. The characteristic property of the characteristic functions $\chi : U \to \mathbb{R}$ is that they are idempotent in the sense that $\chi(u)\chi(u) = \chi(u)$ for all $u \in U$, and the lifted characteristic property of the projection operators $P : V \to V$ is that they are idempotent in the sense that $P^2 : V \to V \to V = P : V \to V$. Finally, the "spectral decomposition" of a set attribute lifts to the spectral decomposition of a *vector space attribute*:

$$f = \textstyle\sum_r r\chi_{f^{-1}(r)} : U \to \mathbb{R} \text{ lifts to } L = \sum_\lambda \lambda P_\lambda : V \to V$$
Lift of a set attribute to a vector space attribute

Thus a vector space attribute is just a linear operator whose eigenspaces span the whole space which is called a *diagonalizable linear operator* [16]. Then we see that the proper lift of a set attribute using the basis principle does indeed define a vector space partition, namely that of the eigenspaces of a diagonalizable linear operator, and that the values of the attribute are constant on the blocks of the vector space partition–as desired. To keep the eigenvalues of the linear operator real, quantum mechanics restricts the vector space attributes to *Hermitian* (or *self-adjoint*) linear operators, which represent *observables*, on a Hilbert space.

Hermann Weyl is one of the few quantum physicists who, in effect, outlined the lifting program connecting QM/sets and QM. He called a partition a "grating" or "sieve," and then considered both set partitions and vector space partitions (direct sum decompositions) as the respective types

[4] There are two ways to think of the "set version" of a concept: as a straight set concept with no mention of vector spaces over $\mathbb{Z}_2$, or as a vector space over $\mathbb{Z}_2$ concept (which already starts to combine set and vector space concepts). For instance, the pure set concept of the partition given by an attribute $f : U \to \mathbb{R}$ is the set partition $\left\{f^{-1}(r)\right\}_r$ and the "direct sum" is the set disjoint union $U = \biguplus_r f^{-1}(r)$. But this can be recast in $\mathbb{Z}_2^{|U|}$ as the vector space direct sum: $\wp(U) = \sum_r \oplus_\wp \left(f^{-1}(r)\right)$ of the vector space partition $\left\{\wp\left(f^{-1}(r)\right)\right\}_r$.

of gratings.[27, pp. 255-257] He started with a numerical attribute on a set, which defined the set partition or "grating" [27, p. 255] with blocks having the same attribute-value. Then he moved to the quantum case where the set or "aggregate of n states has to be replaced by an n-dimensional Euclidean vector space" [27, p. 256] (note the lift from cardinality n sets to dimension n vector spaces). 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" [27, p. 256], i.e., a vector space partition (direct sum decomposition of the space).

Lifting Program	Set concept: QM over sets ($\mathbb{Z}_2$)	Vector concept: QM over $\mathbb{C}$
Eigenvalues	r s.t. $f\restriction S = rS$ for some S	λ s.t. $Lv = \lambda v$ for some v
Eigenvectors	S s.t. $f\restriction S = rS$ for some r	v s.t. $Lv = \lambda v$ for some λ
Eigenspaces	$\{S: f\restriction S = rS\} = \wp(f^{-1}(r))$	$\{v: Lv = \lambda v\} = W_\lambda$
Eigenspace Partition	Set partition of "eigenspaces" $U = \uplus f^{-1}(r)$	Vector space partition of eigenspaces $V = \Sigma \oplus W_\lambda$
Characteristic functions	$\chi_S: U \to \{0,1\}$ for subsets S like $f^{-1}(r)$	Projection operators for subspaces like $W_\lambda = P_\lambda(V)$
Spectral decomposition	Set attribute $f: U \to \mathbb{R}$: $f = \Sigma_r \, r\chi_{f^{-1}(r)}$	Hermitian linear operator: $L = \Sigma_\lambda \, \lambda P_\lambda$

Figure 1: Set numerical attributes lift to linear operators

2.5 Lifting compatible attributes to commuting operators

Since two set attributes $f : U \to \mathbb{R}$ and $g : U' \to \mathbb{R}$ define two inverse image partitions $\{f^{-1}(r)\}$ and $\{g^{-1}(s)\}$ on their domains, we need to extend the concept of compatible partitions to the attributes that define the partitions. That is, two attributes $f : U \to \mathbb{R}$ and $g : U' \to \mathbb{R}$ are *compatible* if they have the same domain $U = U'$. We have previously lifted the notion of compatible set partitions to compatible vector space partitions. Since real-valued set attributes lift to Hermitian linear operators, the notion of compatible set attributes just defined would lift to two linear operators being *compatible* if their eigenspace partitions are compatible. It is a standard fact of QM math (e.g., [17, pp. 102-3] or [16, p. 177]) that two (Hermitian) linear operators $L, M : V \to V$ are compatible if and only if they commute, $LM = ML$. Hence the *commutativity* of linear operators is the lift of the compatibility (i.e., defined on the same set) of set attributes. Thus the join of two eigenspace partitions is defined iff the operators commute. Weyl also pointed this out: "Thus combination [join] of two gratings [vector space partitions] presupposes commutability...". [27, p. 257]

Given two compatible set attributes $f : U \to \mathbb{R}$ and $g : U \to \mathbb{R}$, the join of their "eigenspace" partitions has as blocks the non-empty intersections $f^{-1}(r) \cap g^{-1}(s)$. Each block in the join of the "eigenspace" partitions could be characterized by the ordered pair of "eigenvalues" (r, s). An "eigenvector" of f, $S \subseteq f^{-1}(r)$, and of g, $S \subseteq g^{-1}(s)$, would be a "simultaneous eigenvector": $S \subseteq f^{-1}(r) \cap g^{-1}(s)$.

In the lifted case, two commuting Hermitian linear operator L and M have compatible eigenspace partitions $W_L = \{W_\lambda\}$ (for the eigenvalues λ of L) and $W_M = \{W_\mu\}$ (for the eigenvalues μ of M). The blocks in the join $W_L \vee W_M$ of the two compatible eigenspace partitions are the non-zero

subspaces $\{W_\lambda \cap W_\mu\}$ which can be characterized by the ordered pairs of eigenvalues (λ, μ). The nonzero vectors $v \in W_\lambda \cap W_\mu$ are *simultaneous eigenvectors* for the two commuting operators, and there is a basis for the space consisting of simultaneous eigenvectors.[5]

A set of compatible set attributes is said to be *complete* if the join of their partitions is the discrete partition (the blocks have cardinality 1). Each element of U is then characterized by the ordered n-tuple $(r, ..., s)$ of attribute values.

In the lifted case, a set of commuting linear operators is said to be *complete* if the join of their eigenspace partitions is nondegenerate, i.e., the blocks have dimension 1. The eigenvectors that generate those one-dimensional blocks of the join are characterized by the ordered n-tuples $(\lambda, ..., \mu)$ of eigenvalues so the eigenvectors are usually denoted as the eigenkets $|\lambda, ..., \mu\rangle$ in the Dirac notation. These *Complete Sets of Commuting Operators* are Dirac's CSCOs [8].

2.6 Summary of the QM/sets-to-QM bridge

The lifting program or bridge developed so far is summarized in the following table.

Lifting Summary	Set concept	Vector space concept
Partition	Direct sum decomposition π = {B} of U: U = $\uplus$B	Direct sum decomposition $\{W_i\}$ of V: V = $\sum\oplus W_i$
Real-valued Attribute	Function f:U$\to\mathbb{R}$	Hermitian operator L: V$\to$V
Partition of attribute	Inverse-image partition $\{f^{-1}(r)\}$ for f:U$\to\mathbb{R}$	Eigenspace partition W_L = $\{W_\lambda\}$ for L: V$\to$V
Compatible partitions	Partitions π,σ on same set U	Vector space partitions $\{W_i\}$ and $\{X_j\}$ with common basis
Compatible attributes	Attributes f,g:U$\to\mathbb{R}$ defined on same set U	Commuting operators LM = ML, i.e., common basis of simultaneous eigenvectors.
Join of compatible attribute partitions	$f^{-1}\vee g^{-1} = \{f^{-1}(r)\cap g^{-1}(s)\}$ for f,g:U$\to\mathbb{R}$	$W_L\vee W_M = \{W_\lambda \cap W_\mu\}$ for LM = ML
CSCO	Singleton blocks of $\vee f_i^{-1}$ for compatible attributes $\{f_i^{-1}\}$	One-dim. blocks of $\vee W_{L_i}$ for commuting operators $\{L_i\}$

Figure 2: Summary of Lifting Program

3 The probability calculus in QM/sets

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

The set version of QM is said to be "over $\mathbb{Z}_2$" since the power set $\wp(U)$ (for a finite non-empty universe set U) is a vector space over $\mathbb{Z}_2 = \{0, 1\}$ where the subset addition $S + T$ is the *symmetric difference* (or inequivalence) of subsets, i.e., $S + T = S \not\equiv T = S \cup T - S \cap T$ for $S, T \subseteq U$. Given a finite universe set $U = \{u_1, ..., u_n\}$ of cardinality n, the U-basis in $\mathbb{Z}_2^n$ is the set of singletons

[5] One must be careful not to assume that the simultaneous eigenvectors are the eigenvectors for the operator $LM = ML$ due to the problem of degeneracy.

$\{u_1\}, \{u_2\}, ..., \{u_n\}$ and a vector in $\mathbb{Z}_2^n$ is specified in the U-basis by its $\mathbb{Z}_2$-valued characteristic function $\chi_S : U \to \mathbb{Z}_2$ for an subset $S \subseteq U$ (e.g., a string of n binary numbers). Similarly, a vector v in $\mathbb{C}^n$ is specified in terms of an orthonormal basis $\{|v_i\rangle\}$ by a $\mathbb{C}$-valued function $\langle \_ | v \rangle : \{v_i\} \to \mathbb{C}$ assigning a complex amplitude $\langle v_i | v \rangle$ to each basis vector. One of the key pieces of mathematical machinery in QM, namely the inner product, does not exist in vector spaces over finite fields but basis-dependent "brackets" can still be defined and a norm or absolute value can be defined to play a similar role in the probability algorithm of QM/sets.[6]

Seeing $\wp(U)$ as the vector space $\mathbb{Z}_2^{|U|}$ allows different bases in which the vectors can be expressed (as well as the basis-free notion of a vector as a ket, since only the bra is basis-dependent). Consider the simple case of $U = \{a, b, c\}$ where the U-basis is $\{a\}$, $\{b\}$, and $\{c\}$. But the three subsets $\{a, b\}$, $\{b, c\}$, and $\{a, b, c\}$ also form a basis since: $\{a, b\} + \{a, b, c\} = \{c\}$; $\{b, c\} + \{c\} = \{b\}$; and $\{a, b\} + \{b\} = \{a\}$. These new basis vectors could be considered as the basis-singletons in another equicardinal universe $U' = \{a', b', c'\}$ where $a' = \{a, b\}$, $b' = \{b, c\}$, and $c' = \{a, b, c\}$. In the following *ket table*, each row is a ket of $V = \mathbb{Z}_2^3$ expressed in the U-basis, the U'-basis, and a U''-basis.

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

Vector space isomorphism: $\mathbb{Z}_2^3 \cong \wp(U) \cong \wp(U') \cong \wp(U'')$ where row = ket.

3.2 The brackets

In a Hilbert space, the inner product is used to define the amplitudes $\langle v_i | v \rangle$ and the norm $|v| = \sqrt{\langle v | v \rangle}$, and the probability algorithm can be formulated using this norm. In a vector space over $\mathbb{Z}_2$, the Dirac notation can still be used but in a basis-dependent form (like matrices as opposed to operators) that defines a real-valued norm even though there is no inner product. The kets $|S\rangle$ for $S \subseteq U$ are basis-free but the corresponding bras are basis-dependent. For $u \in U$, the "*bra*" $\langle \{u\} |_U : \wp(U) \to \mathbb{R}$ is defined by the "*bracket*":

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

Then $\langle \{u_i\} |_U \{u_j\} \rangle = \chi_{\{u_j\}}(u_i) = \chi_{\{u_i\}}(u_j) = \delta_{ij}$ is the set-version of $\langle v_i | v_j \rangle = \delta_{ij}$ (for an orthonormal basis $\{|v_i\rangle\}$). Assuming a finite U, the "bracket" linearly extends to the more general basis-dependent form (where $|S|$ is the cardinality of S):

$$\langle T |_U S \rangle = |T \cap S| \text{ for } T, S \subseteq U.^7$$

This basis principle can be run in reverse to "delift" a vector space concept to sets. Consider an orthonormal basis set $\{|v_i\rangle\}$ in a finite dimensional Hilbert space. Given two subsets $T, S \subseteq \{|v_i\rangle\}$ of the basis set, consider the unnormalized superpositions $\psi_T = \sum_{|v_i\rangle \in T} |v_i\rangle$ and $\psi_S = \sum_{|v_i\rangle \in S} |v_i\rangle$. Then their inner product in the Hilbert space is $\langle \psi_T | \psi_S \rangle = |T \cap S|$, which "delifts" (crossing the

[6] Often scare quotes, as in "brackets," are used to indicate the named concept in QM/sets as opposed to full QM–although this may also be clear from the context.

[7] Thus $\langle T |_U S \rangle = |T \cap S|$ takes values outside the base field of $\mathbb{Z}_2$ just like the Hamming distance function $|T + S|$ on vector spaces over $\mathbb{Z}_2$ in coding theory [19, p. 66] as applied to pairs of sets represented as binary strings.

bridge in the other direction) to $\langle T|_U S\rangle = |T \cap S|$ for subsets $T, S \subseteq U$ of the U-basis of $\mathbb{Z}_2^{|U|}$. In both cases, the bracket gives the size of the overlap.

3.3 Ket-bra resolution

The basis-dependent "*ket-bra*" $|\{u\}\rangle \langle\{u\}|_U$ is the "one-dimensional" projection operator:

$$|\{u\}\rangle \langle\{u\}|_U = \{u\} \cap () : \wp(U) \to \wp(U)$$

and the "ket-bra identity" holds as usual:

$$\textstyle\sum_{u\in U} |\{u\}\rangle \langle\{u\}|_U = \sum_{u\in U} (\{u\} \cap ()) = I : \wp(U) \to \wp(U)$$

where the summation is the symmetric difference of sets in $\mathbb{Z}_2^n$. The overlap $\langle T|_U S\rangle$ can be resolved using the "ket-bra identity" in the same basis: $\langle T|_U S\rangle = \sum_u \langle T|_U \{u\}\rangle \langle\{u\}|_U S\rangle$. Similarly a ket $|S\rangle$ can be resolved in the U-basis;

$$|S\rangle = \textstyle\sum_{u\in U} |\{u\}\rangle \langle\{u\}|_U S\rangle = \sum_{u\in U} \langle\{u\}|_U S\rangle |\{u\}\rangle = \sum_{u\in U} |\{u\} \cap S| \{u\}$$

where a subset $S \subseteq U$ is just expressed as the sum of the singletons $\{u\} \subseteq S$. That is ket-bra resolution in sets. 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 basis-dependent bra $\langle\{u\}|_U$ 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 $\langle\{u\}|_U S\rangle = |\{u\} \cap S|$.[8]

3.4 The norm

Then the (basis-dependent) *U-norm* $\|S\|_U : \wp(U) \to \mathbb{R}$ is defined, as usual, as the square root of the bracket:[9]

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

for $S \in \wp(U)$ which is the set-version of the basis-free norm $|\psi| = \sqrt{\langle\psi|\psi\rangle}$ (since the inner product does not depend on the basis). Note that a ket has to be expressed in the U-basis to apply the basis-dependent definition so in the above example, $\|\{a'\}\|_U = \sqrt{2}$ since $\{a'\} = \{a, b\}$ in the U-basis.

3.5 The Born rule

For a specific basis $\{|v_i\rangle\}$ and for any nonzero vector v in a finite dimensional complex vector space, $|v|^2 = \sum_i \langle v_i|v\rangle \langle v_i|v\rangle^*$ (* is complex conjugation) whose set version would be: $\|S\|_U^2 = \sum_{u\in U} \langle\{u\}|_U S\rangle^2$. Since

$$|v\rangle = \textstyle\sum_i \langle v_i|v\rangle |v_i\rangle \text{ and } |S\rangle = \sum_{u\in U} \langle\{u\}|_U S\rangle |\{u\}\rangle,$$

applying the Born rule by squaring the coefficients $\langle v_i|v\rangle$ and $\langle\{u\}|_U S\rangle$ (and normalizing) gives the probabilities of the eigen-elements v_i or $\{u\}$ given a state v or S in QM and QM/sets:

$$\textstyle\sum_i \frac{\langle v_i|v\rangle\langle v_i|v\rangle^*}{|v|^2} = 1 \text{ and } \sum_u \frac{\langle\{u\}|_U S\rangle^2}{\|S\|_U^2} = \sum_u \frac{|\{u\} \cap S|}{|S|} = 1$$

where $\frac{\langle v_i|v\rangle\langle v_i|v\rangle^*}{|v|^2}$ is a 'mysterious' quantum probability while $\frac{\langle\{u\}|_U S\rangle^2}{\|S\|_U^2} = \frac{|\{u\} \cap S|}{|S|}$ is the unmysterious Laplacian equal probability $\Pr(\{u\}|S)$ rule for getting u when sampling S.[10]

[8] The term "$\{u\} \cap S'$" is not even defined since it is the intersection of subsets of two different universes. One of the luxuries of having a basis independent inner product in QM over $\mathbb{C}$ is being able to ignore bases in the bra-ket notation.

[9] We use the double-line notation $\|S\|_U$ for the norm of a set to distinguish it from the single-line notation $|S|$ for the cardinality of a set, whereas the customary absolute value notation for the norm of a vector in full QM is $|v|$.

[10] Note that there is no notion of a normalized vector in a vector space over $\mathbb{Z}_2$ (another consequence of the lack of an inner product). The normalization is, as it were, postponed to the probability algorithm which is computed in the rationals.

3.6 Spectral decomposition on sets

An observable, i.e., a Hermitian operator, on a Hilbert space determines its 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." As previously noted, the connection between the numerical attributes $f : U \to \mathbb{R}$ of QM/sets and the Hermitian operators of QM is established by "seeing" the function f as a formal operator: $f \upharpoonright () : \wp(U) \to \wp(U)$. Applied to the basis elements $\{u\} \subseteq U$, we may write $f \upharpoonright \{u\} = f(u)\{u\} = r\{u\}$ as the set-version of an eigenvalue equation applied to an eigenvector where the multiplication $r\{u\}$ is only formal (read $r\{u\}$ as: the function f takes the value r on $\{u\}$). Then for any subset $S \subseteq f^{-1}(r)$ where f is constant, we may also formally write: $f \upharpoonright S = rS$ as an "eigenvalue equation" satisfied by all the "eigenvectors" S in the "eigenspace" $\wp\left(f^{-1}(r)\right)$, a subspace of $\wp(U)$, for the "eigenvalue" r. Since $f^{-1}(r) \cap () : \wp(U) \to \wp(U)$ is the projection operator[11] to the "eigenspace" $\wp\left(f^{-1}(r)\right)$ for the "eigenvalue" r, we have the spectral decomposition for a Hermitian operator $L = \sum_\lambda \lambda P_\lambda$ in QM and for a U-attribute $f : U \to \mathbb{R}$ in QM/sets:

$$L = \textstyle\sum_\lambda \lambda P_\lambda : V \to V \text{ and } f \upharpoonright () = \sum_r r\left(f^{-1}(r) \cap ()\right) : \wp(U) \to \wp(U)$$

Spectral decomposition of operators in QM and QM/sets.

When the base field increases from $\mathbb{Z}_2$ to $\mathbb{R}$ or $\mathbb{C}$, then the formal multiplication $r\left(f^{-1}(r) \cap ()\right)$ is internalized as an actual multiplication, and the projection operator $f^{-1}(r) \cap ()$ on sets becomes a projection operator on a vector space over $\mathbb{R}$ or $\mathbb{C}$. Thus the operator representation $L = \sum_\lambda \lambda P_\lambda$ of an observable numerical attribute is just the internalization of a numerical attribute made possible by the enriched base field $\mathbb{R}$ or $\mathbb{C}$. Similarly, the set brackets $\langle T|_U S\rangle$ taking values outside the base field $\mathbb{Z}_2$ become internalized as an inner product with the same enrichment of the base field. It is the comparative "poverty" of the base field $\mathbb{Z}_2$ that requires the QM/sets "brackets" to take "de-internalized" or "externalized" values outside the base field and for a formal multiplication to used in the operator presentation $f \upharpoonright () = \sum_r r\left(f^{-1}(r) \cap ()\right)$ of a numerical attribute $f : U \to \mathbb{R}$.[12] Or put the other way around, the only numerical attributes that can be internally represented in $\wp(U) \cong \mathbb{Z}_2^n$ are the characteristic functions $\chi_S : U \to \mathbb{Z}_2$ that are internally represented in the U-basis as the projection operators $S \cap () : \wp(U) \to \wp(U)$.

3.7 Completeness and orthogonality of projection operators

The usual completeness and orthogonality conditions on eigenspaces also have set-versions in QM over $\mathbb{Z}_2$:

1. completeness: $\sum_\lambda P_\lambda = I : V \to V$ has the set-version: $\sum_r f^{-1}(r) \cap () = I : \wp(U) \to \wp(U)$, and
2. orthogonality: for $\lambda \neq \lambda'$, $P_\lambda P_{\lambda'} = 0 : V \to V$ (where 0 is the zero operator) has the set-version: for $r \neq r'$, $\left[f^{-1}(r) \cap ()\right]\left[f^{-1}(r') \cap ()\right] = \emptyset \cap () : \wp(U) \to \wp(U)$.[13]

[11] Since $\wp(U)$ is now interpreted as a vector space, it should be noted that the projection operator $T \cap () : \wp(U) \to \wp(U)$ is not only idempotent but linear, i.e., $(T \cap S_1) + (T \cap S_2) = T \cap (S_1 + S_2)$. Indeed, this is the distributive law when $\wp(U)$ is interpreted as a Boolean ring.

[12] In the engineering literature, eigenvalues are seen as "stretching or shrinking factors" but that is *not* their role in QM. The whole machinery of eigenvectors [e.g., $f \upharpoonright \{u\} = r\{u\}$], eigenspaces [e.g., $\wp\left(f^{-1}(r)\right)$], and eigenvalues [e.g., $f(u) = r$] in QM is a way of representing a numerical attribute [e.g., $f : U \to \mathbb{R}$ in the set case] *inside* a vector space that has a rich enough base field.

[13] 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 their overlap (instead of their inner product) being 0.

3.8 Measuring attributes on sets

The Pythagorean results (for the complete and orthogonal projection operators):

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

give the probabilities for measuring attributes. Since

$$|S| = \|S\|_U^2 = \sum_r \left\|f^{-1}(r) \cap S\right\|_U^2 = \sum_r \left|f^{-1}(r) \cap S\right|$$

we have in QM and in QM/sets:

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

where $\frac{|P_\lambda(v)|^2}{|v|^2}$ is the quantum probability of getting λ in an L-measurement of v while $\frac{\left|f^{-1}(r)\cap S\right|}{|S|}$ has the rather unmysterious interpretation of the probability $\Pr(r|S)$ of the random variable $f : U \to \mathbb{R}$ having the "eigen-value" r when sampling $S \subseteq U$. Thus the set-version of the Born rule is not some weird "quantum" notion of probability on sets but the perfectly ordinary Laplace-Boole rule for the conditional probability $\frac{\left|f^{-1}(r)\cap S\right|}{|S|}$, given $S \subseteq U$, of a random variable $f : U \to \mathbb{R}$ having the value r.

3.9 Contextuality

Given a ket $|S\rangle$, the probability of getting another ket $|\{a\}\rangle$ as an outcome of a measurement in QM/sets will depend on the context in terms of the measurement basis. In the previous ket table, comparing sets in the U-basis and U''-basis, we see that $\{a,b\} = \{b''\}$ (or in the ket notation: $|\{a,b\}\rangle = |\{b''\}\rangle$) and $\{a\} = \{a''\}$. Taking $S = \{a,b\}$, the probability of getting $\{a\}$ in a U-basis measurement is: $\Pr(\{a\}|S) = |\{a\}\cap\{a,b\}|/|\{a,b\}| = 1/2$. But taking the same ket $|\{a,b\}\rangle = |\{b''\}\rangle$ as the given state and measuring in the U''-basis, the probability of getting the ket $|\{a\}\rangle = |\{a''\}\rangle$ is: $\Pr(\{a''\}|\{b''\}) = |\{a''\}\cap\{b''\}|/|\{b''\}| = 0$.

3.10 The objective indefiniteness interpretation

On top of the mathematics of QM/sets, there is an objective indefiniteness interpretation which is just the set-version of the objective indefiniteness interpretation of QM developed elsewhere [11]. The collecting-together of some elements $u \in U$ into a subset $S \subseteq U$ is interpreted as the superposition of the "eigen-elements" $u \in S$ to form an "indefinite element" S (with the vector sum $S = \sum_{u\in U} \langle\{u\}|_U S\rangle \{u\}$ in the vector space $\wp(U)$ over $\mathbb{Z}_2$ giving the superposition).[14]

The indefinite element S is being "measured" using the "observable" f where the probability $\Pr(r|S)$ of getting the "eigenvalue" r is $\frac{\left|f^{-1}(r)\cap S\right|}{|S|}$ and where the "damned quantum jump" goes from S 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. That state represents a more-definite element $f^{-1}(r) \cap S$ that now has the definite f-value of r–so a second measurement would yield the same "eigenvalue" r and the same vector $f^{-1}(r) \cap \left[f^{-1}(r) \cap S\right] = f^{-1}(r) \cap S$ using the idempotency of the set-version of projection operators (all as in the standard Dirac-von-Neumann treatment of measurement). These questions of interpretation will not be emphasized here where the focus is on the mathematical relationship between QM/sets and full QM.

[14] In logic, a *choice function* is a function $\varepsilon()$ that applied to a non-empty subset $S \subseteq U$ picks out an element $\varepsilon(S) = u \in S$ (or equivalently a singleton $\varepsilon(S) = \{u\} \subseteq S$). The indeterminancy of a choice function is, as it were, where stochasticity enters QM. For finite sets, we might consider a probabilistic choice function that would pick out any element (or singleton) of S with the equal probability $1/|S|$. A (non-degenerate) "measurement" in QM/sets is a "physical" version of a probabilistic choice function; it goes from an indefinite entity S to some definite entity $\{u\} \subseteq S$ with the probability $1/|S|$.

3.11 Summary of the probability calculus

These set-versions and more (the average value of an attribute is treated later) are summarized in the following table for a finite U and a finite dimensional Hilbert space V with $\{|v_i\rangle\}$ as any orthonormal basis.

Vector space over $\mathbb{Z}_2$: QM/sets	Hilbert space case: QM over $\mathbb{C}$
Projections: $S \cap () : \wp(U) \to \wp(U)$	$P : V \to V$
Spectral Decomp.: $f \restriction () = \sum_r r\left(f^{-1}(r) \cap ()\right)$	$L = \sum_\lambda \lambda P_\lambda$
Compl.: $\sum_r f^{-1}(r) \cap () = I : \wp(U) \to \wp(U)$	$\sum_\lambda P_\lambda = I$
Orthog.: $r \neq r'$, $\left[f^{-1}(r) \cap ()\right]\left[f^{-1}(r') \cap ()\right] = \emptyset \cap ()$	$\lambda \neq \lambda'$, $P_\lambda P_{\lambda'} = 0$
Brackets: $\langle S|_U T\rangle = \lvert S \cap T\rvert$ = overlap for $S, T \subseteq U$	$\langle\psi|\varphi\rangle$ = "overlap" of ψ and φ
Ket-bra: $\sum_{u\in U} \lvert\{u\}\rangle\langle\{u\}|_U = \sum_{u\in U}(\{u\} \cap ()) = I$	$\sum_i \lvert v_i\rangle\langle v_i\rvert = I$
Resolution: $\langle S|_U T\rangle = \sum_u \langle S|_U \{u\}\rangle\langle\{u\}|_U T\rangle$	$\langle\psi|\varphi\rangle = \sum_i \langle\psi|v_i\rangle\langle v_i|\varphi\rangle$
Norm: $\lVert S\rVert_U = \sqrt{\langle S|_U S\rangle} = \sqrt{\lvert S\rvert}$ where $S \subseteq U$	$\lvert\psi\rvert = \sqrt{\langle\psi|\psi\rangle}$
Pythagoras: $\lVert S\rVert_U^2 = \sum_{u\in U}\langle\{u\}|_U S\rangle^2 = \lvert S\rvert$	$\lvert\psi\rvert^2 = \sum_i \langle v_i|\psi\rangle^* \langle v_i|\psi\rangle$
Laplace: $S \neq \emptyset$, $\sum_{u\in U} \frac{\langle\{u\}|_U S\rangle^2}{\lVert S\rVert_U^2} = \sum_{u\in S} \frac{1}{\lvert S\rvert} = 1$	$\lvert\psi\rangle \neq 0$, $\sum_i \frac{\langle v_i|\psi\rangle^*\langle v_i|\psi\rangle}{\lvert\psi\rvert^2} = \frac{\lvert\langle v_i|\psi\rangle\rvert^2}{\lvert\psi\rvert^2} = 1$
Born: $\lvert S\rangle = \sum_{u\in U}\langle\{u\}|_U S\rangle\lvert\{u\}\rangle$, $\Pr(u|S) = \frac{\langle\{u\}|_U S\rangle^2}{\lVert S\rVert_U^2}$	$\lvert\psi\rangle = \sum_i \langle v_i|\psi\rangle\lvert v_i\rangle$, $\Pr(v_i|\psi) = \frac{\lvert\langle v_i|\psi\rangle\rvert^2}{\lvert\psi\rvert^2}$
$\lVert S\rVert_U^2 = \sum_r \left\lVert f^{-1}(r) \cap S\right\rVert_U^2 = \sum_r \left\lvert f^{-1}(r) \cap S\right\rvert = \lvert S\rvert$	$\lvert\psi\rvert^2 = \sum_\lambda \lvert P_\lambda(\psi)\rvert^2$
$S \neq \emptyset$, $\sum_r \frac{\left\lVert f^{-1}(r)\cap S\right\rVert_U^2}{\lVert S\rVert_U^2} = \sum_r \frac{\left\lvert f^{-1}(r)\cap S\right\rvert}{\lvert S\rvert} = 1$	$\lvert\psi\rangle \neq 0$, $\sum_\lambda \frac{\lvert P_\lambda(\psi)\rvert^2}{\lvert\psi\rvert^2} = 1$
Measurement: $\Pr(r|S) = \frac{\left\lVert f^{-1}(r)\cap S\right\rVert_U^2}{\lVert S\rVert_U^2} = \frac{\left\lvert f^{-1}(r)\cap S\right\rvert}{\lvert S\rvert}$	$\Pr(\lambda|\psi) = \frac{\lvert P_\lambda(\psi)\rvert^2}{\lvert\psi\rvert^2}$
Average of attribute: $\langle f\rangle_S = \frac{\langle S|_U f\restriction()|S\rangle}{\langle S|_U S\rangle}$	$\langle L\rangle_\psi = \frac{\langle\psi|L|\psi\rangle}{\langle\psi|\psi\rangle}$.

Probability mathematics for QM over $\mathbb{Z}_2$ and for QM over $\mathbb{C}$

4 Measurement in QM/sets

4.1 Measurement as partition join operation

In QM/sets, numerical attributes $f : U \to \mathbb{R}$ can be considered as equiprobable random variables on a set of outcomes U. The inverse images of attributes (or random variables) define set partitions $\left\{f^{-1}(r)\right\}$ on the set of outcomes 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*: $\pi \vee \sigma$ 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 $\operatorname{dit}(\pi) \subseteq U \times U$ on U where the ordered pairs (u, u') in $\operatorname{dit}(\pi)$ are the *distinctions* or *dits* of π in the sense that u and u' are in distinct blocks of π. These dit sets $\operatorname{dit}(\pi)$ as binary relations might be called "partition relations" but they are also the "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, $\operatorname{indit}(\pi)$, as a binary relation is just the equivalence relation associated with the partition π.

In the duality between the ordinary Boolean logic of subsets (usually mis-specified as "propositional" logic) and the logic of partitions ([10] or [12]), the elements of a subset and the distinctions of a partition are dual concepts. The partial ordering of subsets in the powerset Boolean algebra $\wp(U)$ is the inclusion of elements and the refinement ordering of partitions on U is just the inclusion of dit sets, i.e., $\sigma \preceq \pi$ iff $\operatorname{dit}(\sigma) \subseteq \operatorname{dit}(\pi)$. The partial ordering in each case is a lattice where the top of the Boolean lattice is the subset U of all possible elements and the top of the lattice of

partitions 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 makes no distinctions.

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

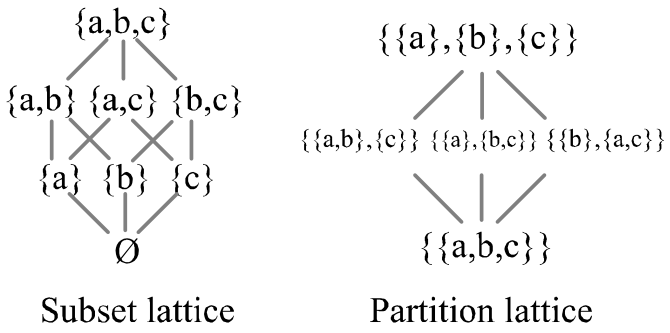

Figure 3: Subset and partition lattices

In the correspondences between QM/sets and QM, a block in a partition on U [i.e., a vector in $\wp(U)$] corresponds to pure state in QM (a state vector in a quantum state space), and a partition on U can be thought of as a mixture of orthogonal pure states with the probabilities given by the probability calculus on QM/sets. Given a "pure state" $S \subseteq U$, the possible results of a non-degenerate U-measurement are the blocks of the discrete partition $\{\{u\}\}_{u \in S}$ on S with each singleton being equiprobable. Each such measurement would have one of the potential "eigenstates" $\{u\} \subseteq S$ as the actual result.

In QM, measurements make distinctions that turn a pure state into a mixture. The abstract essentials of measurement are represented in QM/sets as a distinction-creating processes of turning a "pure state" S into a "mixed state" partition on S (with "distinctions" as defined above in partition logic). The distinction-creating process of "measurement" in QM/sets is the partition join of the indiscrete partition $\{S\}$ (taking S as the universe) and the inverse-image partition $\left\{f^{-1}(r)\right\}$ of the numerical attribute $f : U \to \mathbb{R}$ restricted to S. Again Weyl gets it right. Weyl refers to a partition as a "grating" or "sieve" and then notes that "Measurement means application of a sieve or grating" [27, p. 259], e.g., the application (i.e., join) of the set-grating $\left\{f^{-1}(r)\right\}_r$ to the "pure state" $\{S\}$ to give the "mixed state" $\left\{S \cap f^{-1}(r)\right\}_r$.

4.2 Nondegenerate measurements

In the simple example illustrated below, we start at the one block or "state" of the indiscrete partition or blob which is the completely indistinct element $\{a, b, c\}$. A measurement always 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 indefinite element $\{a, b, c\}$ (since there are four partitions that refine the blob).

For an example of a "nondegenerate measurement," consider any attribute $f : U \to \mathbb{R}$ which has the discrete partition as its inverse image, such as the ordinal number of a letter in the alphabet: $f(a) = 1$, $f(b) = 2$, and $f(c) = 3$. This attribute or "observable" has three "eigenvectors": $f \upharpoonright \{a\} = 1\{a\}$, $f \upharpoonright \{b\} = 2\{b\}$, and $f \upharpoonright \{c\} = 3\{c\}$ with the corresponding "eigenvalues." The "eigenvectors" are $\{a\}$, $\{b\}$, and $\{c\}$, the blocks in the discrete partition of U. Starting in the "pure state" $S = \{a, b, c\}$, a U-measurement using the observable f gives the "mixed state":

$$\{U\} \vee \left\{f^{-1}(r)\right\}_{r=1,2,3} = \mathbf{0} \vee \mathbf{1} = \mathbf{1}.$$

Each such measurement would return an "eigenvalue" r with the probability of $\Pr(r|S) = \frac{|f^{-1}(r) \cap S|}{|S|} = \frac{1}{3}$.

A "projective measurement" makes distinctions in the measured "state" that are sufficient to induce the "quantum jump" or "projection" to the "eigenvector" associated with the observed "eigenvalue." If the observed "eigenvalue" was 3, then the "state" $\{a, b, c\}$ "projects" to $f^{-1}(3) \cap \{a, b, c\} = \{c\} \cap \{a, b, c\} = \{c\}$ as pictured below.

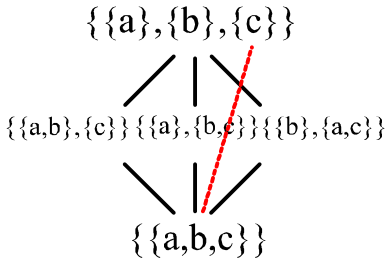

Figure 4: "Nondegenerate measurement"

It might be emphasized that this is an objective state reduction (or "collapse of the wave packet") from the single indefinite element $\{a, b, c\}$ to the single definite element $\{c\}$, not a subjective removal of ignorance as if the "state" had all along been $\{c\}$. For instance, Pascual Jordan in 1934 argued that:

> the electron is forced to a decision. We compel it to assume a definite position; previously, in general, it was neither here nor there; it had not yet made its decision for a definite position... . ... [W]e ourselves produce the results of the measurement. (quoted in [18, p. 161])

This might be illustrated using Weyl's notion of a partition as a "sieve or grating" [27, p. 259] that is applied in a measurement. We might think of a grating as a series of regular polygonal shapes that might be imposed on an indefinite blob of dough. In a measurement, the blob of dough falls through one of the polygonal holes with equal probability and then takes on that shape.

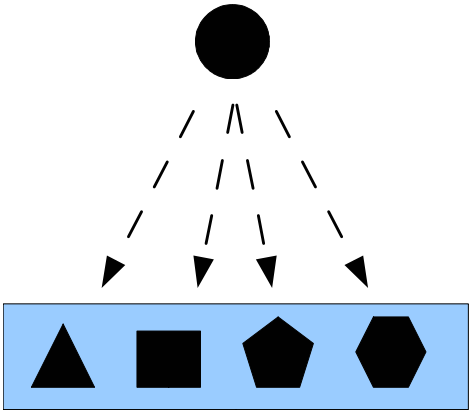

Figure 5: Measurement as randomly giving an indefinite blob of dough a regular polygonal shape.

4.3 Degenerate measurements

For an example of a "degenerate measurement," we choose an attribute with a non-discrete inverse-image partition such as $\pi = \{\{a\}, \{b, c\}\}$. 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 one of the two "eigenspaces" is not a singleton of an eigen-element, the "eigenvalue" of 1 is a set version of a "degenerate eigenvalue." This attribute $\chi_{\{b,c\}}$ has four (non-zero) "eigenvectors": $\chi_{\{b,c\}} \restriction \{b, c\} = 1\{b, c\}$, $\chi_{\{b,c\}} \restriction \{b\} = 1\{b\}$, $\chi_{\{b,c\}} \restriction \{c\} = 1\{c\}$, and $\chi_{\{b,c\}} \restriction \{a\} = 0\{a\}$.

The "measuring apparatus" makes distinctions by "joining" the "observable" partition

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

with the "pure state" which is the single block representing the indefinite element $S = U = \{a, b, c\}$. A measurement apparatus of that "observable" returns one of "eigenvalues" with certain probabilities:

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

Suppose it returns the "eigenvalue" 1. Then the indefinite element $\{a, b, c\}$ "jumps" to the "projection" $\chi^{-1}_{\{b,c\}}(1) \cap \{a, b, c\} = \{b, c\}$ of the "state" $\{a, b, c\}$ to that "eigenvector" [6, p. 221].

Since this is a "degenerate" result (i.e., the "eigenspaces" don't all have "dimension" one), another measurement is needed to make more distinctions. Measurements by attributes that give either of the other two partitions, $\{\{a, b\}, \{c\}\}$ or $\{\{b\}, \{a, c\}\}$, suffice to distinguish $\{b, c\}$ into $\{b\}$ or $\{c\}$, so either attribute together with the attribute $\chi_{\{b,c\}}$ would form a *complete set of compatible attributes* (i.e., the set version of a CSCO). The join of the two attributes' partitions gives the discrete partition. Taking 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 "observables": $\{a\} = |0, 1\rangle$, $\{b\} = |1, 1\rangle$, and $\{c\} = |1, 0\rangle$ (using Dirac's ket-notation to give the ordered pairs).

The second "projective measurement" of the indefinite "superposition" element $\{b, c\}$ using the attribute $\chi_{\{a,b\}}$ with the "eigenspace" partition $\chi^{-1}_{\{a,b\}} = \{\{a, b\}, \{c\}\}$ would induce a jump to either $\{b\}$ 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.

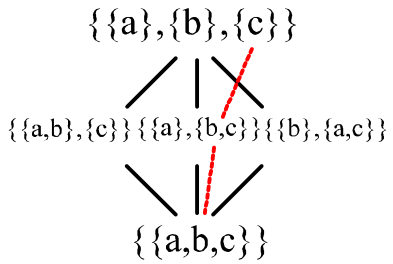

Figure 6: "Degenerate measurement"

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

5 "Time" evolution in QM/sets

The different "de-internalized" treatment of the "brackets" in QM/sets gives a probability calculus, unlike Schumacher and Westmoreland's "modal quantum theory." [21] But both theories agree that evolution of the quantum states over $\mathbb{Z}_2$ is given by non-singular linear transformations. These transformations are, of course, reversible like the unitary transformations of full QM but "unitary" is not defined in the absence of an inner product. QM/sets nevertheless has basis-dependent "brackets" and those "brackets" are preserved if we change the basis along with the non-singular transformation. Let $A : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ be a non-singular transformation where the images of the U-basis $A\,|\{u\}\rangle$ are taken as the basis vectors $\{u'\}$ of a U'-basis. Then for $S, T \subseteq U$, we have the following preservation of the "brackets":

$$\langle T|_U S\rangle = \langle AT|_{AU} AS\rangle = \langle T'|_{U'} S'\rangle$$

where $AT = T' \subseteq U' = AU$ and $AS = S' \subseteq U' = AU$.

In the objective indefiniteness interpretation of QM based on partition logic [11], von Neumann's type 1 processes (measurements) and type 2 processes (unitary evolution) [26] are modeled respectively as the processes that make distinctions or that don't make any distinctions in the strong sense of preserving the degree of indistinctness $\langle\varphi|\psi\rangle$ between quantum states. That characterization of evolution carries over to QM/sets since it is precisely the non-singular transformations that preserve distinctness of QM/sets quantum states, i.e., distinctness of non-zero vectors in $\mathbb{Z}_2^n$.

By rendering QM concepts in the simple context of sets, QM/sets gives an understanding of the basic logic of the QM concept. Much effort has been expended in the philosophy of QM to understand measurement. We have seen that by rendering QM measurement in the context of sets that it is the distinction-making process of applying the partition $\left\{f^{-1}(r)\right\}_r$ of an observable attribute to a pure state partition $\{S\}$ (i.e., taking the partition join) to get the mixed state partition $\left\{f^{-1}(r) \cap S\right\}_r$. Now we see that time evolution in QM (i.e., a degree-of-indistinctness $\langle\psi|\varphi\rangle$ preserving transformation) is modeled in QM/sets by distinction-preserving non-singular transformations. This explains von Neumann's classification of the two types of quantum processes: the distinction-making or type 1 processes (measurement) and the distinction-preserving or type 2 processes (time evolution). In this manner, QM/sets brings out the essence or "logic" of the full QM concepts of measurement and time evolution.

6 Interference without "waves" in QM/sets

The role of the so-called "waves" in ordinary quantum mechanics can be further clarified by viewing quantum dynamics in QM/sets. In QM over $\mathbb{C}$, suppose the Hamiltonian H has an orthonormal basis of energy eigenstate $\{|E_j\rangle\}$. Then the application of the unitary propagation operator $U(t)$ from $t = 0$ to time t applied to $|\psi_0\rangle = \sum_j c_j |E_j\rangle$ has the action:

$$U(t)\,|\psi_0\rangle = |\psi_t\rangle = e^{iHt}\,|\psi_0\rangle = \textstyle\sum_j c_j e^{iHt}\,|E_j\rangle = \sum_j c_j e^{iE_j t}\,|E_j\rangle.$$

Thus $U(t)$ transforms the orthonormal basis $\{|E_j\rangle\}$ into the orthonormal basis $\left\{\left|E'_j\right\rangle\right\} = \left\{e^{iE_j t}\,|E_j\rangle\right\}$.[15] Even though this unitary transformation introduces different relative phases for the different energy eigenstates in $U(t)\,|\psi_0\rangle$, the probabilities for an energy measurement do not change since

[15] Indeed, a *unitary* operator on an inner product space can be defined as a linear operator that transforms an orthonormal basis into an orthonormal basis.

$|c_j|^2 = \left|c_j e^{iE_j t}\right|^2$. The effects of time evolution show when the evolved state $U(t)|\psi_0\rangle$ is measured in *another* basis $\{|a_k\rangle\}$. Suppose for each j, $|E_j\rangle = \sum_k \alpha_k^j |a_k\rangle$ so that:

$$U(t)|\psi_0\rangle = |\psi_t\rangle = \sum_j c_j e^{iE_j t}|E_j\rangle = \sum_j c_j e^{iE_j t}\sum_k \alpha_k^j |a_k\rangle = \sum_k \left(\sum_j c_j e^{iE_j t}\alpha_k^j\right)|a_k\rangle.$$

Then under time evolution, there is interference in the coefficient $\sum_j c_j e^{iE_j t}\alpha_k^j$ of each eigenstate $|a_k\rangle$. Since the complex exponentials $e^{iE_j t}$ can be mathematically interpreted as "waves," this is the interference characteristic of wave-like behavior in the evolution of the quantum state $|\psi_0\rangle$.

But there is interference without waves in QM/sets where many of the characteristic phenomena of QM can nevertheless be reproduced (see later sections on the two-slit experiment and Bell's Theorem). Suppose we start with a state $S \subseteq U = \{u_1, ..., u_n\}$ which is represented in the U-basis as $|S\rangle = \sum_j \langle u_j|_U S\rangle |u_j\rangle = \sum_j b_j |u_j\rangle$ where $\langle u_j|_U S\rangle = b_j \in \mathbb{Z}_2$.[16] Then the "dynamics" of a nonsingular transformation $A : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ takes the basis $\{|u_j\rangle\}$ to another basis $\left\{\left|u'_j\right\rangle\right\}$ (where $A|u_j\rangle = \left|u'_j\right\rangle$) which is the set or binary vector space version of $U(t)$ taking the orthonormal basis $\{|E_j\rangle\}$ to the orthonormal basis $\left\{\left|E'_j\right\rangle\right\}$ where $\left|E'_j\right\rangle = e^{iE_j t}|E_j\rangle$. Thus $|S\rangle$ is transformed, by linearity, into $|S'\rangle = \sum_j b_j \left|u'_j\right\rangle$ with the same b_j's so that $\Pr(u_j|S) = \frac{b_j^2}{|S|} = \frac{b_j^2}{|S'|} = \Pr\left(u'_j|S'\right)$ and $\langle S|_U T\rangle = \langle S'|_{U'} T'\rangle$ (where for $T \subseteq U$, $A|T\rangle = |T'\rangle$ for some $T' \subseteq U'$). But the state $|S'\rangle = \sum_j b_j \left|u'_j\right\rangle$ could be measured in another U''-basis $\left\{\left|u''_j\right\rangle\right\}$ where $\left|u'_j\right\rangle = \sum_k \alpha_k^j \left|u''_k\right\rangle$ so that:

$$A|S\rangle = |S'\rangle = \sum_j b_j \left|u'_j\right\rangle = \sum_j b_j \sum_k \alpha_k^j \left|u''_k\right\rangle = \sum_k \left(\sum_j b_j \alpha_k^j\right)\left|u''_j\right\rangle.$$

Then under time evolution, there is interference in the coefficient $\sum_j b_j \alpha_k^j$ of each eigenstate $\left|u''_j\right\rangle$. This suffices to give the interference phenomena that are ordinarily seen as characteristic of wave-like behavior but there is not even the mathematics of waves in QM/sets. The mathematics of waves (complex exponentials $e^{i\varphi}$) comes into the mathematics of quantum mechanics *only over* $\mathbb{C}$; real exponentials either grow or decay but don't behave as waves.

The following table summarizes the results using the minimal superpositions: $|S\rangle = b_1|u_1\rangle + b_2|u_2\rangle$ and $|\psi_0\rangle = c_1|E_1\rangle + c_2|E_2\rangle$.

QM/sets	QM
$\|u_j\rangle \xrightarrow{A} \left\|u'_j\right\rangle$	$\|E_j\rangle \xrightarrow{U} \left\|E'_j\right\rangle = e^{ig_j t}\|E_j\rangle$
$\|S\rangle = b_1\|u_1\rangle + b_2\|u_2\rangle \to b_1\|u'_1\rangle + b_2\|u'_2\rangle$	$\|\psi_0\rangle = c_1\|E_1\rangle + c_2\|E_2\rangle \to c_1\|E'_1\rangle + c_2\|E'_2\rangle$
$\left\|u'_j\right\rangle = \sum_k \langle u''_k\|_{U''} u'_j\rangle \left\|u''_k\right\rangle = \sum_k \alpha_k^j \left\|u''_k\right\rangle$	$\|E_j\rangle = \sum_k \alpha_k^j \|a_k\rangle$; $\left\|E'_j\right\rangle = e^{ig_j t}\sum_k \alpha_k^j \|a_k\rangle$
$b_1\|u_1\rangle + b_2\|u_2\rangle \to \sum_k \left(b_1\alpha_k^1 + b_2\alpha_k^2\right)\left\|u''_k\right\rangle$	$c_1\|E_1\rangle + c_2\|E_2\rangle \to \sum_k \left(c_1 e^{ig_1 t}\alpha_k^1 + c_2 e^{ig_2 t}\alpha_k^2\right)\|a_k\rangle$

Table showing the role in interference in QM/sets and in QM

Thus QM/sets allows us to tease the QM behavior due to interference apart from the specifically wave-version of that interference in QM over $\mathbb{C}$. The root of the interference is superposition, i.e., the different j's in the coefficients $\sum_j c_j e^{iE_j t}\alpha_k^j$ in QM or $\sum_j b_j \alpha_k^j$ in QM/sets, and superposition is the mathematical representation of indefiniteness. It is indefiniteness that is the basic feature, and a particle in a superposition state for a certain observable will have the evolution of that indefiniteness expressed by coefficients $\sum_j c_j e^{iE_j t}\alpha_k^j$ using complex exponentials (i.e., the mathematics of waves) so the indefiniteness will then appear as "wave-like" behavior–even though the so-called "wave functions" $|\psi_0\rangle$ of QM do not represent physical waves.

[16] For notational simplicity, we will often leave the curly brackets off the singletons $\{u_j\}$ and just write $|u_j\rangle$ instead of $|\{u_j\}\rangle$.

7 Double-slit experiment in QM/sets

QM/sets represents the logical essence of full QM without any of the physical assumptions. Hence to delift the double-slit experiment to QM/sets, we need to imagine the elements of some U-basis as "positions" and a non-singular matrix A as giving the dynamic evolution for one "time" period.

Consider the dynamics given in terms of the U-basis where: $\{a\} \to \{a,b\}$; $\{b\} \to \{a,b,c\}$; and $\{c\} \to \{b,c\}$ in one time period. This is represented by the non-singular one-period change of state matrix:

$$A = \begin{bmatrix} \langle \{a\} \,|_U \{a,b\} \rangle & \langle \{a\} \,|_U \{a,b,c\} \rangle & \langle \{a\} \,|_U \{b,c\} \rangle \\ \langle \{b\} \,|_U \{a,b\} \rangle & \langle \{b\} \,|_U \{a,b,c\} \rangle & \langle \{b\} \,|_U \{b,c\} \rangle \\ \langle \{c\} \,|_U \{a,b\} \rangle & \langle \{c\} \,|_U \{a,b,c\} \rangle & \langle \{c\} \,|_U \{b,c\} \rangle \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

If we take the U-basis vectors as "vertical position" eigenstates, we can device a QM/sets version of the double-slit experiment which models "all of the mystery of quantum mechanics" [13, 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.

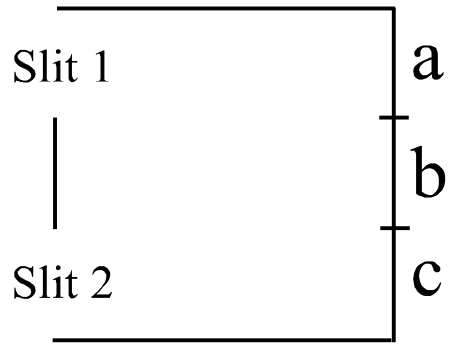

Figure 7: Two-slit setup

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:

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

If the particle was measured at slit 1, i.e., was in the post-measurement 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:

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

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:

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

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:

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

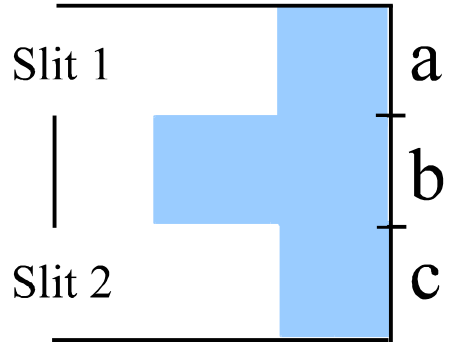

Figure 8: Final probability distribution with measurement at the slits.

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 $\{a,b\}+\langle b,c\rangle = \{a,c\}$ 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:

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

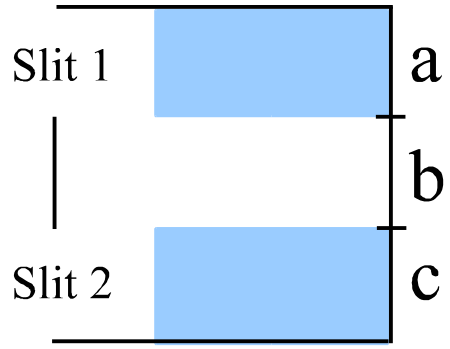

Figure 9: Final probability distribution with no measurement at slits

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 $\{a, b\} + \langle b, c\rangle = \{a, c\}$. The addition is, of course, mod 2 (where $-1 = +1$) so, in "wave language," the two "wave crests" 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-measurement yielding either of those eigenstates. Figure 9 shows the simplest example of the "light and dark bands" characteristic of superposition and interference illustrating "all of the mystery of quantum mechanics".

This model gives the simplest logical essence of the two-slit experiment without the complex-valued wave functions that distract from the essential point; 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 without "waves".

8 Entanglement in QM/sets

A QM concept that generates much interest is entanglement. Hence it might be useful to consider "entanglement" in QM/sets.

First we need to establish the connections across the set-vector-space bridge by lifting the set notion of the direct (or Cartesian) product $X \times Y$ of two sets X and Y. Using the basis principle, we apply the set concept to the two basis sets $\{v_1, ..., v_m\}$ and $\{w_1, ..., w_n\}$ of two vector spaces V and W (over the same base field) and then we see what it generates. The set direct product of the two basis sets is the set of all ordered pairs (v_i, w_j), which we will write as $v_i \otimes w_j$, and then we generate the vector space, denoted $V \otimes W$, over the same base field from those basis elements $v_i \otimes w_j$. That vector space is the *tensor product*, and it is not in general the direct product $V \times W$ of the vector spaces. The cardinality of $X \times Y$ is the product of the cardinalities of the two sets, and the dimension of the tensor product $V \otimes W$ is the product of the dimensions of the two spaces (while the dimension of the direct product $V \times W$ is the sum of the two dimensions).

A vector $z \in V \otimes W$ is said to be *separated* if there are vectors $v \in V$ and $w \in W$ such that $z = v \otimes w$; otherwise, z is said to be *entangled.* Since vectors delift to subsets, a subset $S \subseteq X \times Y$ is said to be *separated* or a *product* if there exists subsets $S_X \subseteq X$ and $S_Y \subseteq Y$ such that $S = S_X \times S_Y$; otherwise $S \subseteq X \times Y$ is said to be *entangled.* In general, let S_X be the support or projection of S on X, i.e., $S_X = \{x : \exists y \in Y, (x, y) \in S\}$ and similarly for S_Y. Then S is separated iff $S = S_X \times S_Y$.

For any subset $S \subseteq X \times Y$, where X and Y are finite sets, a natural measure of its "entanglement" can be constructed by first viewing S as the support of the equiprobable or Laplacian joint probability distribution on S. If $|S| = N$, then define $\Pr(x, y) = \frac{1}{N}$ if $(x, y) \in S$ and $\Pr(x, y) = 0$ otherwise.

The marginal distributions[17] are defined in the usual way:

$$\Pr(x) = \sum_y \Pr(x, y)$$
$$\Pr(y) = \sum_x \Pr(x, y).$$

A joint probability distribution $\Pr(x, y)$ on $X \times Y$ is *independent* if for all $(x, y) \in X \times Y$,

$$\Pr(x, y) = \Pr(x) \Pr(y).$$
Independent distribution

Otherwise $\Pr(x, y)$ is said to be *correlated.*

Proposition 1 *A subset $S \subseteq X \times Y$ is "entangled" iff the equiprobable distribution on S is correlated (non-independent).*

[17] The marginal distributions are the set versions of the reduced density matrices of QM.

Proof: If S is "separated", i.e., $S = S_X \times S_Y$, then $\Pr(x) = |S_Y|/N$ for $x \in S_X$ and $\Pr(y) = |S_X|/N$ for $y \in S_Y$ where $|S_X||S_Y| = N$. Then for $(x, y) \in S$,

$$\Pr(x, y) = \frac{1}{N} = \frac{N}{N^2} = \frac{|S_X||S_Y|}{N^2} = \Pr(x)\Pr(y)$$

and $\Pr(x, y) = 0 = \Pr(x)\Pr(y)$ for $(x, y) \notin S$ so the equiprobable distribution is independent. If S is "entangled," i.e., $S \neq S_X \times S_Y$, then $S \subsetneqq S_X \times S_Y$ so let $(x, y) \in S_X \times S_Y - S$. Then $\Pr(x), \Pr(y) > 0$ but $\Pr(x, y) = 0$ so it is not independent, i.e., is correlated. □

Consider the set version of one qubit space where $U = \{a, b\}$. The product set $U \times U$ has 15 nonempty subsets. Each factor U of $U \times U$ has 3 nonempty subsets so $3 \times 3 = 9$ of the 15 subsets are separated subsets leaving 6 entangled subsets.

$S \subseteq U \times U$
$\{(a, a), (b, b)\}$
$\{(a, b), (b, a)\}$
$\{(a, a), (a, b), (b, a)\}$
$\{(a, a), (a, b), (b, b)\}$
$\{(a, b), (b, a), (b, b)\}$
$\{(a, a), (b, a), (b, b)\}$

The six entangled subsets

The first two are the "Bell states" which are the two graphs of bijections $U \longleftrightarrow U$ and have the maximum entanglement if entanglement is measured by the logical divergence $d(\Pr(x, y)||\Pr(x)\Pr(y))$[9]. All the 9 separated states have zero entanglement by the same measure.

For an entangled subset S, a sampling x of left-hand system will change the probability distribution for a sampling of the right-hand system y, $\Pr(y|x) \neq \Pr(y)$. In the case of maximal "entanglement" (e.g., the "Bell states"), when S is the graph of a bijection between U and U, the value of y is *determined* by the value of x (and vice-versa).

9 Bell's Theorem in QM/sets

A simple version of a Bell inequality can be derived in the case of $\mathbb{Z}_2^2$ where the only three bases are: $U = \{a, b\}$, $U' = \{a', b'\}$, and $U'' = \{a'', b''\}$, with the relations given in the ket table:

kets	U-basis	U'-basis	U''-basis
$\|1\rangle$	$\{a, b\}$	$\{a'\}$	$\{a''\}$
$\|2\rangle$	$\{b\}$	$\{b'\}$	$\{a'', b''\}$
$\|3\rangle$	$\{a\}$	$\{a', b'\}$	$\{b''\}$
$\|4\rangle$	$\emptyset$	$\emptyset$	$\emptyset$

Ket table for $\wp(U) \cong \wp(U') \cong \wp(U'') \cong \mathbb{Z}_2^2$.

Attributes defined on the three universe sets U, U', and U'', such as say $\chi_{\{a\}}$, $\chi_{\{b'\}}$, and $\chi_{\{a''\}}$, are incompatible as can be seen in several ways. For instance the set partitions defined on U and U', namely $\{\{a\}, \{b\}\}$ and $\{\{a'\}, \{b'\}\}$, cannot be obtained as two different ways to partition the same set since $\{a\} = \{a', b'\}$ and $\{a'\} = \{a, b\}$, i.e., an "eigenstate" in one basis is a superposition in the other. The same holds in the other pairwise comparison of U and U'' and of U' and U''.

Given a ket in $\mathbb{Z}_2^2 \cong \wp(U) \cong \wp(U') \cong \wp(U'')$, 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	b	a'	b'	a''	b''
$\{a,b\}=\{a'\}=\{a''\}$	$\frac{1}{2}$	$\frac{1}{2}$	1	0	1	0
$\{b\}=\{b'\}=\{a'',b''\}$	0	1	0	1	$\frac{1}{2}$	$\frac{1}{2}$
$\{a\}=\{a',b'\}=\{b''\}$	1	0	$\frac{1}{2}$	$\frac{1}{2}$	0	1

State-outcome probability table.

The delift of the tensor product of vector spaces is the Cartesian or direct product of sets, and the delift of the vectors in the tensor product are the subsets of direct product of sets (as seen in the above treatment of entanglement in QM/sets). Thus in the U-basis, the basis elements are the elements of $U\times U$ and the "vectors" are all the subsets in $\wp(U\times U)$. But we could obtain the same "space" as $\wp(U'\times U')$ and $\wp(U''\times U'')$, 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:

$$\mathbb{Z}_2^2\otimes\mathbb{Z}_2^2\cong\wp(U\times U)\cong\wp(U'\times U')\cong\wp(U''\times U'').$$

Since $\{a\}=\{a',b'\}=\{b''\}$ and $\{b\}=\{b'\}=\{a'',b''\}$, the subset $\{a\}\times\{b\}=\{(a,b)\}\subseteq U\times U$ is expressed in the $U'\times U'$-basis as $\{a',b'\}\times\{b'\}=\{(a',b'),(b',b')\}$, and in the $U''\times U''$-basis it is $\{b''\}\times\{a'',b''\}=\{(b'',a''),(b'',b'')\}$. Hence one row in the ket table has:

$$\{(a,b)\}=\{(a',b'),(b',b')\}=\{(b'',a''),(b'',b'')\}.$$

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

$U\times U$	$U'\times U'$	$U''\times U''$
$\{(a,a)\}$	$\{(a',a'),(a',b'),(b',a'),(b',b')\}$	$\{(b'',b'')\}$
$\{(a,b)\}$	$\{(a',b'),(b',b')\}$	$\{(b'',a''),(b'',b'')\}$
$\{(b,a)\}$	$\{(b',a'),(b',b')\}$	$\{(a'',b''),(b'',b'')\}$
$\{(b,b)\}$	$\{(b',b')\}$	$\{(a'',a''),(a'',b''),(b'',a''),(b'',b'')\}$
$\{(a,a),(a,b)\}$	$\{(a',a'),(b',a')\}$	$\{(b'',a'')\}$
$\{(a,a),(b,a)\}$	$\{(a',a'),(a',b')\}$	$\{(a'',b'')\}$
$\{(a,a),(b,b)\}$	$\{(a',a'),(a',b'),(b',a')\}$	$\{(a'',a''),(a'',b''),(b'',a'')\}$
$\{(a,b),(b,a)\}$	$\{(a',b'),(b',a')\}$	$\{(a'',b''),(b'',a'')\}$

Partial ket table for $\wp(U\times U)\cong\wp(U'\times U')\cong\wp(U''\times U'')$

As before, we can classify each subset as separated or entangled and we can furthermore see how that is independent of the basis. For instance $\{(a,a),(a,b)\}$ is separated since:

$$\{(a,a),(a,b)\}=\{a\}\times\{a,b\}=\{(a',a'),(b',a')\}=\{a',b'\}\times\{a'\}=\{(b'',a'')\}=\{b''\}\times\{a''\}.$$

An example of an entangled state is:

$$\{(a,a),(b,b)\}=\{(a',a'),(a',b'),(b',a')\}=\{(a'',a''),(a'',b''),(b'',a'')\}.$$

Taking this entangled state as the initial state, the probability of getting the state $\{a\}$ by performing a U-basis measurement on the left-hand system is:

$$\Pr(\{(a,-)\}\,|\,\{(a,a),(b,b)\})=\tfrac{|\{(a,a)\}|}{|\{(a,a),(b,b)\}|}=\tfrac{1}{2}.$$[18]

[18] In full QM, performing a measurement of an operator A on the left-hand system is interpreted as performing an $A\otimes I$ measurement on the whole system. In QM/sets, $\mathbb{Z}_2^2\times\mathbb{Z}_2^2$ is spanned by the $U\times U$-basis but also by the $U\times U'$-basis and the $U\times U''$-basis. If $f:U\to 2$ is an attribute on U and $g:X\to 2$ is an attribute on X where X could be U,U', or U'', then $f\times g$ is defined on the $U\times X$-basis by $f\times g((u,x))=f(u)\,g(x)$. If $g=1$ (constant function 1 on

The probability of getting the state $\{a'\}$ by performing a U'-basis measurement on the left-hand system is:

$$\Pr\left(\{(a',-)\}\,|\,\{(a',a'),(a',b'),(b',a')\}\right)=\frac{\left|\{(a',a'),(a',b')\}\right|}{|\{(a',a'),(a',b'),(b',a')\}|}=\frac{2}{3}.$$

The probability of getting the state $\{a''\}$ by performing a U''-basis measurement on the left-hand system is:

$$\Pr\left(\{(a'',-)\}\,|\,\{(a'',a''),(a'',b''),(b'',a'')\}\right)=\frac{\left|\{(a'',a''),(a'',b'')\}\right|}{|\{(a'',a''),(a'',b''),(b'',a'')\}|}=\frac{2}{3}.$$

The probability of each of these outcomes occurring (if each is done instead of either of the others) is the product of the conditional probabilities. Then there is a probability distribution on $U\times U'\times U''$, all conditionalized by the same entangled state, where:

$$\begin{gathered}\Pr(a,a',a'')\\=\Pr\left(\{(a,-)\}\,|\,\{(a,a),(b,b)\}\right)\\\times\Pr\left(\{(a',-)\}\,|\,\{(a',a'),(a',b'),(b',a')\}\right)\\\times\Pr\left(\{(a'',-)\}\,|\,\{(a'',a''),(a'',b''),(b'',a'')\}\right)\\=\frac{1}{2}\frac{2}{3}\frac{2}{3}=\frac{2}{9}.\end{gathered}$$

In this way, a probability distribution $\Pr(x,y,z)$ is defined on $U\times U'\times U''$.

A Bell inequality can be obtained from this joint probability distribution over the outcomes $U\times U'\times U''$ of measuring these three incompatible attributes [7]. Consider the following marginals:

$$\begin{aligned}\Pr(a,a')&=\Pr(a,a',a'')+\Pr(a,a',b'')\checkmark\\\Pr(b',b'')&=\Pr(a,b',b'')\checkmark+\Pr(b,b',b'')\\\Pr(a,b'')&=\Pr(a,a',b'')\checkmark+\Pr(a,b',b'')\checkmark.\end{aligned}$$

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:

$$\Pr(a,a')+\Pr(b',b'')\geq\Pr(a,b'')$$
Bell inequality.

X), then $f\times 1_X[(u,x)]=f(u)$ on the three bases $U\times U$, $U\times U'$, and $U\times U''$. Thus $\left|(f\times 1_X)^{-1}(r)\right|=\left|f^{-1}(r)\times X\right|$ and hence the unconditional probabilities for a left-measurement in the U-basis are unambiguous:

$$\Pr(r)=\frac{|(f\times 1_X)^{-1}(r)|}{|U\times X|}$$

for any X. But ambiguity arises in the conditional probabilities for a given ket $|S\rangle$ since the size of set representing the ket may differ between bases. In the case at hand, $S=\{(a,a),(b,b)\}\subseteq U\times U$ but the same ket is represented by $S'=\{(a,a'),(a,b'),(b,b')\}\subseteq U\times U'$ in the $U\times U'$-basis. If $f=\chi_{\{a\}}:U\to 2$, then the left-measurement for $\{a\}$, i.e., for the eigenvalue 1, is ambiguous depending on the basis $U\times U$ or $U\times U'$ for the entire product space $\mathbb{Z}_2^2\times\mathbb{Z}_2^2$. In the first case, $\frac{|(f\times 1_U)^{-1}(1)\cap S|}{|S|}=\frac{1}{2}$ and in the second case, $\frac{\left|(f\times 1_{U'})^{-1}(1)\cap S'\right|}{|S'|}=\frac{2}{3}$. This ambiguity in the notion of a left or right measurement on a product is due ultimately to the basis-dependent brackets in QM/sets. Unlike full QM which has basis-independent brackets, QM/sets violates parameter independence [23] since choosing X is choosing a "parameter" in QM/sets (thanks to Jerry Finkelstein for raising this question about QM/sets).

For the purpose of deriving a Bell inequality, we adopt the convention of always interpreting a left or right measurement in a certain basis X as assuming the $X\times X$-basis on the product. Under this convention, Bell's inequality will still be derived, and then shown to be violated in QM/sets.

All this has to do with measurements on the left-hand system. But the "Bell state" is left-right symmetrical so the same probabilities would be obtained if we used a right-hand system measurement:

$\Pr(\{(a,-)\} \mid \{(a,a),(b,b)\}) = \Pr(\{(-,a)\} \mid \{(a,a),(b,b)\}) = \frac{1}{2}$;
$\Pr(\{(b,-)\} \mid \{(a,a),(b,b)\}) = \Pr(\{(-,b)\} \mid \{(a,a),(b,b)\}) = \frac{1}{2}$;
$\Pr(\{(a',-)\} \mid \{(a',a'),(a',b'),(b',a')\}) = \Pr(\{(-,a')\} \mid \{(a',a'),(a',b'),(b',a')\}) = \frac{2}{3}$;
$\Pr(\{(b',-)\} \mid \{(a',a'),(a',b'),(b',a')\}) = \Pr(\{(-,b')\} \mid \{(a',a'),(a',b'),(b',a')\}) = \frac{1}{3}$;
$\Pr(\{(a'',-)\} \mid \{(a'',a''),(a'',b''),(b'',a'')\}) = \Pr(\{(-,a'')\} \mid \{(a'',a''),(a'',b''),(b'',a'')\}) = \frac{2}{3}$;
and
$\Pr(\{(b'',-)\} \mid \{(a'',a''),(a'',b''),(b'',a'')\}) = \Pr(\{(-,b'')\} \mid \{(a'',a''),(a'',b''),(b'',a'')\}) = \frac{1}{3}$.[19]
This is analogous to the assumption that each sock in a pair of socks will have the same properties.[1, Chap. 16] Hence the right-hand measurements give the same probability distribution and the same inequality.

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 $\Pr(a,a')$ is the probability of a U-measurement on the left-hand system giving $\{a\}$ *and then in addition* (not instead of) a U'-measurement on the right-hand system giving $\{a'\}$, and so forth.

This is a crucial step in the argument so it worth being very clear using subscripts.

- Step 1: $\Pr(a,a')_1$ is the probability of getting $\{a\}$ in a left U-measurement and getting $\{a'\}$ *if instead* a left U'-measurement was made so:

$$\Pr(a,a')_1 = \Pr(\{(a,-)\} \mid \{(a,a),(b,b)\}) \times \Pr(\{(a',-)\} \mid \{(a',a'),(a',b'),(b',a')\}) = \tfrac{1}{2}\tfrac{2}{3} = \tfrac{1}{3}.$$

- Step 2: $\Pr(a,a')_2$ is the probability of getting $\{a\}$ in a left U-measurement and getting $\{a'\}$ if instead a *right* U'-measurement was made so:

$$\Pr(a,a')_2 = \Pr(\{(a,-)\} \mid \{(a,a),(b,b)\}) \times \Pr(\{(-,a')\} \mid \{(a',a'),(a',b'),(b',a')\}) = \tfrac{1}{2}\tfrac{2}{3} = \tfrac{1}{3}.$$

- Step 3: $\Pr(a,a')_3$ is the probability of getting $\{a\}$ in a left U-measurement and, *under the assumption of independence of the left-right measurements*, *also* (not instead of) getting $\{a'\}$ in a right U'-measurement:

$$\Pr(a,a')_3 = \Pr(\{(a,-)\} \mid \{(a,a),(b,b)\}) \times \Pr(\{(-,a')\} \mid \{(a',a'),(a',b'),(b',a')\}) = \tfrac{1}{2}\tfrac{2}{3} = \tfrac{1}{3}.$$

Hence the joint probability distribution would be the same and the above Bell inequality:

$$\Pr(a,a')_3 + \Pr(b',b'')_3 \geq \Pr(a,b'')_3$$

would still hold *under the independence assumption* using the step 3 probabilities in all cases. 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 the conditional probabilities for the given entangled state.

To compute $\Pr(a,a')_3$, we first measure the left-hand component in the U-basis. Since $\{(a,a),(b,b)\}$ is the given state, and (a,a) and (b,b) 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 $\{a'\}$ (i.e., "eigenvalue" 0 for the "observable" $\chi_{\{b'\}}$) is $\frac{1}{2}$ (as seen in the state-outcome table). Thus the probability is $\Pr(a,a')_3 = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$.

To compute $\Pr(b',b'')_3$, we first perform a U'-basis "measurement" on the left-hand component of the given state $\{(a,a),(b,b)\} = \{(a',a'),(a',b'),(b',a')\}$, and we see that the probability of

[19] The same holds for the other "Bell state": $\{(a,b),(b,a)\}$.

getting $\{b'\}$ is $\frac{1}{3}$. Then the right-hand system is in the state $\{a'\}$ and the probability of getting $\{b''\}$ in a U''-basis "measurement" of the right-hand system in the state $\{a'\}$ is 0 (as seen from the state-outcome table). Hence the probability is $\Pr(b', b'')_3 = 0$.

Finally we compute $\Pr(a, b'')_3$ by first making a U-measurement on the left-hand component of the given state $\{(a, a), (b, b)\}$ and get the result $\{a\}$ with probability $\frac{1}{2}$. Then the state of the second system is $\{a\}$ so a U''-measurement will give the $\{b''\}$ result with probability 1 so the probability is $\Pr(a, b'')_3 = \frac{1}{2}$.

Then we plug the probabilities into the Bell inequality:

$$\Pr(a, a')_3 + \Pr(b', b'')_3 \geq \Pr(a, b'')_3$$
$$\tfrac{1}{4} + 0 \not\geq \tfrac{1}{2}$$
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. This result is somewhat less striking in QM/sets than in full QM since QM/sets just shows the bare logic of the Bell argument in the simplest space $\mathbb{Z}_2^2 \otimes \mathbb{Z}_2^2$ without any dramatic physical assumption like a space-like separation between the left-hand and right-hand physical systems.

Part II
Quantum information and computation theory in QM/sets

10 Quantum information theory in QM/sets

10.1 Logical entropy

Obtaining quantum information theory for QM/sets is not a simple matter of delifting the ordinary quantum information theory (QIT). This is because much of QIT is obtained by transporting over or lifting the notion of Shannon entropy from classical information theory (which is then renamed "von Neumann entropy"). Shannon entropy is a higher-level concept adapted for questions of coding and communication; it is not a basic logical concept. Classical information theory itself needs to be refounded on a logical basis using the logical notion of entropy that arises naturally out of partition logic (that is dual to the usual Boolean subset logic). That logical information theory can then be simply reformulated using delifted machinery from QM, namely density matrices, and thus logical information theory is reformulated as "quantum" information theory for QM/sets.

The process is quite analogous to the way that classical logical finite probability was reformulated as the probability calculus for QM/sets. Conceptually, the next step beyond subset logic was the quantitative treatment that gave logical finite probability theory. Historically, Boole presented logical finite probability theory as this quantitative step beyond subset logic in his book entitled: *An Investigation of the Laws of Thought on which are founded the Mathematical Theories of Logic and Probabilities*. The universe U was the finite number of possible outcomes and the subsets were events. Quoting Poisson, Boole defined "the measure of the probability of an event [as] the ratio of the number of cases favourable to that event, to the total number of cases favourable and unfavourable, and all equally possible." [4, p. 253]

Hence one obvious next quantitative step beyond partition logic is to make the analogous conceptual moves and to see what theory emerges. The theory that emerges is a logical version of information theory.

For a finite U, the finite (Laplacian) *probability* $\Pr(S)$ of a subset ("event") is the normalized counting measure on the subset: $\Pr(S) = |S| / |U|$. Analogously, the finite *logical entropy* $h(\pi)$ of a partition π is the normalized counting measure of its dit set: $h(\pi) = |\mathrm{dit}(\pi)| / |U \times U|$. If U is an urn with each "ball" in the urn being equiprobable, then $\Pr(S)$ is the probability of an element randomly drawn from the urn is an element in S, and, similarly, $h(\pi)$ is the probability that a pair of elements randomly drawn from the urn (with replacement) is a distinction of π.

Let $\pi = \{B_1, ..., B_m\}$ with $p_i = |B_i| / |U|$ being the probability of drawing an element of the block B_i. The number of indistinctions (non-distinctions) of π is $|\mathrm{indit}(\pi)| = \Sigma_i |B_i|^2$ so the number of distinctions is $|\mathrm{dit}(\pi)| = |U|^2 - \Sigma_i |B_i|^2$ and thus since $\Sigma_i p_i = 1$, the logical entropy of π is: $h(\pi) = \left[|U|^2 - \Sigma_i |B_i|^2\right] / |U|^2 = 1 - \Sigma_i p_i^2 = (\Sigma_i p_i) - \Sigma_i p_i^2 = \Sigma_i p_i (1 - p_i)$, so that:

$$\text{Logical entropy: } h(\pi) = \Sigma_i p_i (1 - p_i).$$

Shannon's notion of entropy is a high-level notion adapted to communications theory [22]. The Shannon entropy $H(\pi)$ of the partition π (with the same probabilities assigned to the blocks) is:

$$\text{Shannon entropy: } H(\pi) = \Sigma_i p_i \log(1/p_i)$$

where the log is base 2.

Each entropy can be seen as the probabilistic average of the "block entropies" $h(B_i) = 1 - p_i$ and $H(B_i) = \log(1/p_i)$. To interpret the block entropies, consider a special case where $p_i = 1/2^n$ and every block is the same so there are 2^n equal blocks like B_i in the partition. The logical entropy of that special equal-block partition, $\Sigma_i p_i (1 - p_i) = (2^n) p_i (1 - p_i) = (2^n)(1/2^n)(1 - p_i) = 1 - p_i$, is the:

$$\text{Logical block entropy: } h(B_i) = 1 - p_i.$$

Instead of directly counting the distinctions, we could take the number of binary equal-blocked partitions it takes to distinguish all the 2^n blocks in that same partition. As in the game of "twenty questions," if there is a search for an unknown designated block, then each such binary question can reduce the number of blocks by a power of 2 so the minimum number of binary partitions it takes to distinguish all the 2^n blocks (and find the hidden block no matter where it was) is $n = \log(2^n) = \log(1/p_i)$, which is the:

$$\text{Shannon block entropy: } H(B_i) = \log(1/p_i).$$

To precisely relate the block entropies, we solve each for p_i which is then eliminated to obtain:

$$h(B) = 1 - \left(1/2^{H(B)}\right).$$

Exact relation between Shannon and logical block entropies

The interpretation of the Shannon block entropy is then extended by analogy to the general case where $1/p_i$ is not a power of 2 so that the Shannon entropy $H(\pi) = \Sigma_i p_i H(B_i)$ is then interpreted as the *average* number of binary partitions needed to make all the distinctions between the blocks of π—whereas the logical entropy is still the *exact* normalized count $h(\pi) = \Sigma_i p_i h(B_i) = |\mathrm{dit}(\pi)| / |U \times U|$ of the distinctions of the partition π.

The two notions of entropy boil down to two different ways to count the distinctions of a partition. Thus the concept of a distinction from partition logic provides a logical basis for the notion of entropy in information theory.[20]

[20] For further development of logical information theory, see Ellerman [9].

10.2 Density matrices in QM/sets

The notion of logical entropy generalizes naturally to quantum information theory where it also provides a new foundational notion of entropy based on the idea of *information as distinctions* that are preserved in unitary transformations and made objectively in measurements.[11] Our purpose here is to formulate logical entropy using the delifted notion of density matrices which gives QIT/sets, and which then foreshadows how the "classical" logical information theory can be lifted to give a new foundation for the full QIT. The previous treatment of measurement in QM/sets can also be reformulated using density matrices and logical entropy.

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 mimic the usual procedure for forming the density matrix $\rho(\pi)$ for the "orthogonal pure states" B with the probabilities p_B. The (normalized) "pure state" B is represented by the column vector $|B\rangle = \left[\sqrt{q_1}, \sqrt{q_2}, ..., \sqrt{q_n}\right]^t$ where $q_j = 1/|B|$ if $u_j \in B$, and $q_j = 0$ otherwise. Then the *density matrix* $\rho(B)$ *for the pure state* $B \subseteq U$ is then (calculating in the reals):

$$\rho(B) = |B\rangle (|B\rangle)^t = \begin{bmatrix} \sqrt{q_1} \\ \sqrt{q_2} \\ \vdots \\ \sqrt{q_n} \end{bmatrix} \left[\sqrt{q_1}, \sqrt{q_2}, ..., \sqrt{q_n}\right] = \begin{bmatrix} q_1 & \sqrt{q_1 q_2} & \cdots & \sqrt{q_1 q_n} \\ \sqrt{q_2 q_1} & q_2 & \cdots & \sqrt{q_2 q_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{q_n q_1} & \sqrt{q_n q_2} & \cdots & q_n \end{bmatrix}.$$

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

$$\rho(\{a, b\}) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \rho(\{c\}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

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:

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

While this construction mimics the usual construction of the density matrix for orthogonal pure states, the remarkable thing is that the entries have a direct interpretation in terms of the dits and indits of the partition π:

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

All the entries are real "amplitudes" whose squares are the two-draw probabilities of drawing a pair of elements from U (with replacement) that is an indistinction of π. To foreshadow the quantum case, the non-zero entries $\rho_{jk}(\pi) = \sqrt{\frac{1}{|U|}\frac{1}{|U|}} = \frac{1}{|U|}$ indicate that u_j and u_k "cohere" together in a block or "pure state" of the partition, i.e., are an indit of the partition. 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 $\frac{1}{|U|}$. After interchanging some rows and the corresponding columns, the density matrix $\rho(\pi)$ would be a block-diagonal matrix with the blocks corresponding to the blocks B of the partition π.

The *quantum logical entropy* of a density matrix ρ in full QM is: $h(\rho) = 1 - \text{tr}\left[\rho^2\right]$, and the logical entropy of a set partition π with equiprobable points is $h(\pi) = 1 - \sum_{B \in \pi} p_B^2$. The following proposition shows that the above defined density matrix $\rho(\pi)$ in QM/sets was the right definition.

Proposition 2 $h(\pi) = 1 - \text{tr}\left[\rho(\pi)^2\right]$.

Proof: The proof is simplified if we assume that rows and columns have been interchanged so that $\rho(\pi)$ is a block-diagonal matrix with the submatrix-blocks corresponding to the blocks of partition π. If $u_i \in B \in \pi$,then the i^{th} diagonal element of the squared matrix $\rho(\pi)^2$ is $\frac{1}{|U|}\frac{1}{|U|} + ... + \frac{1}{|U|}\frac{1}{|U|}$ ($|B|$ times) or$|B|\left(\frac{1}{|U|}\right)^2$ and that diagonal element will occur $|B|$ times. Hence the trace (sum of diagonal elements) is:

$$\text{tr}\left[\rho(\pi)^2\right] = \sum_{B\in\pi} |B| \times |B| \frac{1}{|U|^2} = \sum_{B\in\pi}\left(\frac{|B|}{|U|}\right)^2 = \sum_{B\in\pi} p_B^2$$

so $h(\pi) = 1 - \sum_B p_B^2$ equals the delifted quantum version: $h(\rho(\pi)) = 1 - \text{tr}\left[\rho(\pi)^2\right]$. □

The logical entropy $h(\pi)$ of a partition is interpreted as the total two-draw probability of drawing a distinction of the partition π. Hence by the above proposition, $\text{tr}\left[\rho(\pi)^2\right]$ is the total probability of drawing an indistinction of π. For a pure state, we have the logical entropy $h(\rho(B)) = 1 - \text{tr}\left[\rho(B)^2\right] = 0$ since the sum of the indistinction probabilities $\text{tr}\left[\rho(B)^2\right]$ is 1 (all pairs are indistinctions in a pure state) while in the general "mixed state" of a partition π (with "orthogonal pure state" blocks $B \in \pi$), $\text{tr}\left[\rho(\pi)^2\right]$ is the sum of the indistinction probabilities.

All this carries over from QM/sets to full QM where it provides *an interpretation of the entries in a density matrix*. Let $\rho = \sum_{i=1}^m \lambda_i |\psi_i\rangle\langle\psi_i|$ be an $n \times n$ density matrix in its orthogonal decomposition so the non-negative eigenvalues λ_i sum to one and the eigenvectors ψ_i are orthonormal. Let $\{|j\rangle : j = 1, ..., n\}$ be an orthonormal eigenvector basis for the whole space so that $\psi_i = \sum_j \alpha_{ij} |j\rangle$ and $\sum_j \alpha_{ij}\alpha_{ij}^* = 1$ where both sums can be taken as only over the j such that $|j\rangle$ has the eigenvalue λ_i (since $\alpha_{ij} = 0$ elsewhere). Previously the square $\rho_{jk}(\pi)^2$ was the two-draw probability for the ordered pair of indices (j, k) if they are in the same block, i.e., are indits of π, otherwise $\rho_{jk}(\pi) = 0$. Similarly, the absolute square $\rho_{jk}\rho_{jk}^*$ of that j, k entry of ρ is nonzero only if $|j\rangle$ and $|k\rangle$ are in the same pure state ψ_i so those probabilities can be interpreted as the *coherence probabilities* for $(|j\rangle, |k\rangle)$ cohering together in the same pure state ψ_i. That is,

$$\rho_{jk}\rho_{jk}^* = \lambda_i\alpha_{ij}\alpha_{ik}^*\lambda_i\alpha_{ij}^*\alpha_{ik} = \rho_{jj}\rho_{kk}$$

which is the probability of getting the ordered pair of eigenvectors $(|j\rangle, |k\rangle)$ in a pair of independent nondegenerate measurements in the $\{|j\rangle\}$ basis–if $|j\rangle$ and $|k\rangle$ cohere together in the same pure state ψ_i. Thus in full QM, $\text{tr}\left[\rho^2\right]$ is the *total coherence probability* while the logical entropy $h(\rho) = 1 - \text{tr}\left[\rho^2\right]$ is the *total decoherence probability*. For a pure state, there are no distinctions or decoherence, so the logical entropy is 0 in both cases. The following table then summarizes the lifting-delifting relationship between the density matrix $\rho(\pi)$ of a partition in QM/sets and (the orthogonal decomposition presentation of) a density matrix ρ in QM.

Density matrix: $\rho(\pi)$ in QM over sets	$\rho = \sum_i \lambda_i \|\psi_i\rangle\langle\psi_i\|$ in QM over $\mathbb{C}$
Disjoint blocks: $B \in \pi$	Orthogonal eigenvectors: $\|\psi_i\rangle$
Block probabilities: $p_B = \frac{\|B\|}{\|U\|}$	Eigenvalues of ρ: λ_i
Point probabilities: $\frac{1}{\|U\|}$	$\lambda_i\alpha_{ij}\alpha_{ij}^* = \rho_{jj}$
Pure state matrix: $\rho(B) = \|B\rangle\langle B\|$	$\rho(\psi_i) = \|\psi_i\rangle\langle\psi_i\|$
Density matrix: $\rho(\pi) = \sum_{B\in\pi} p_B\rho(B)$	$\rho = \sum_i \lambda_i\rho(\psi_i)$
Prob. (j, k) if indit of π: $\rho_{jk}(\pi)^2 = 1/\|U\|^2$	Coherence prob.: $\rho_{jk}\rho_{jk}^* = \rho_{jj}\rho_{kk}$
Logical entropy: $h(\rho(\pi)) = 1 - \text{tr}\left[\rho(\pi)^2\right]$	$h(\rho) = 1 - \text{tr}\left[\rho^2\right]$
$h(\rho(\pi))$ = total distinction probability	$h(\rho)$ = total decoherence prob.
Pure state: $h(\rho(B_i)) = 0$ (no dits)	$h(\rho(\psi_i)) = 0$ (no decoherence)

Density matrices in QM/sets and in QM

Previously we formulated a probability calculus for QM/sets and then noted that it was just the usual logical finite probability theory (in a "non-commutative" version) so that reflects back to give a better understanding of the usual probability calculus in full QM. Now we have formulated the notion of density matrices in QM/sets, and then we noted that it was just a reformulation of logical information theory using the density matrix formalism. Then that reflects back to full QM so that we can now provide an interpretation of the off-diagonal entries in a density matrix ρ as coherence probabilities (like the indistinction probabilities in the set case). And then the quantum logical entropy is the total decoherence probability.

10.3 Density matrices and expectations

Given an attribute $f : U = \{u_1, ..., u_n\} \to \mathbb{R}$, the matrix representing this attribute in QM/sets is:

$$f = \begin{bmatrix} f(1) & 0 & \cdots & 0 \\ 0 & f(2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(n) \end{bmatrix}.$$

Given a subset $S \subseteq U$, the "density matrix" for that state has, with some column and row interchanges, a constant $|S| \times |S|$ block with the values $1/|S|$ and zeros elsewhere:

$$\rho(S) = \begin{bmatrix} \frac{1}{|S|} & \cdots & \frac{1}{|S|} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{|S|} & \cdots & \frac{1}{|S|} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Then, as in full QM, we have the result that the average value of an operator f in a state given by a density matrix $\rho(S)$ is the trace of the product $f\rho(S)$:

$$\begin{aligned} \mathrm{tr}\left[f\rho(S)\right] &= \tfrac{1}{|S|}\sum_{u\in S} f(u) = \tfrac{1}{|S|}\sum_{u\in U} f(u)\,\langle S|_U\,\{u\}\rangle\,\langle\{u\}|_U S\rangle \\ &= \tfrac{1}{|S|}\,\langle S|_U f\restriction()|S\rangle = \tfrac{\langle S|_U f\restriction()|S\rangle}{\langle S|_U S\rangle} = \langle f\rangle_S \end{aligned}$$

where $f\restriction|\{u\}\rangle = f(u)\,|\{u\}\rangle$ and $\sum_u |\{u\}\rangle\langle\{u\}|_U = I$.

10.4 Measuring measurement in QM/sets

A real-valued "observable" is a set attribute $f : U \to \mathbb{R}$ which defines an inverse-image partition $\{f^{-1}(r)\}$. Recall from the logic of partitions that the blocks of the join $\pi \vee \sigma$ of two partitions $\pi = \{B\}$ and $\sigma = \{C\}$ are the non-empty intersections $B \cap C$. This action of the join operation could be considered as a set of projection operators $\{B \cap ()\}_{B\in\pi}$ acting on the blocks $C \in \sigma$–or on a single subset $S \subseteq U$. The partition $f^{-1} = \{f^{-1}(r)\}$ acts as a set of projection operators $f^{-1} \vee () = \{f^{-1}(r) \cap ()\}$ on the "pure-state" S to partition it into the parts $f^{-1} \vee (S) = \{f^{-1}(r) \cap S\}$.

What is the "law of motion" to describe the change in the density matrix resulting from a measurement? Given the density matrix $\rho(S)$ of the "pure state" S, the density matrix $\hat{\rho}(S)$ resulting from the measurement of the observable f is the "mixed state" density matrix $\rho(\pi)$ for the partition given by the join operation $\pi = f^{-1} \vee (S)$. Thus the "law of motion" is the join operation on partitions. That is the canonical way that distinctions are made to move to a more refined partition.

Let's put the previous measurement of the state $S = U$ using the non-degenerate attribute $f(a) = 1, f(b) = 2$, and $f(c) = 3$ in this form using density matrices. The pre-measurement density matrix is the previous $\rho(U)$, the constant matrix with all entries $1/3$. The three projection operators to the eigenspaces of the f-attribute in the U -basis are now:

$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, P_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ and } P_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

instead of $\left\{f^{-1}(r) \cap ()\right\}_{r=1,2,3}$ in the non-matrix version. Hence the "density matrix" for the projection to the eigenspace for $\lambda = 1$ is obtained by first projecting the state $P_1 |U\rangle$ (like $f^{-1}(1) \cap (U) = \{a\}$ in the non-matrix version) and then forming the "density matrix"

$$(P_1 |U\rangle)(P_1 |U\rangle)^t = P_1 \rho(U) P_1 = P_1 \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} P_1$$

$$= P_1 \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} P_1 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P_1 = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so doing the same for the other eigenvalues and summing gives the mixed state density matrix $\hat{\rho}(U)$ that results from the measurement:

$$\hat{\rho}(U) = \sum_{i=1}^{3} P_i \rho(U) P_i = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.$$

The *main result* is that this standard diagonal density matrix representing the result of a non-degenerate measurement is the density matrix $\rho(\pi)$ of the partition formed by the join-action:

$$\pi = f^{-1} \vee (U) = \{\{a\}, \{b\}, \{c\}\} \vee \{a, b, c\} = \{\{a\}, \{b\}, \{c\}\} = \mathbf{1}.$$

Since it was a non-degenerate measurement, all the distinctions were made so all the off-diagonal terms are 0. Each of the off-diagonal terms was "decohered" by the nondegenerate measurement so the post-measurement "amplitude" of (i, j) still "cohering" is 0. The density matrix version of the

$$\rho(U) \overset{measurement}{\longrightarrow} \hat{\rho}(U) = \rho\left(f^{-1} \vee (U)\right)$$
Measurement as join-action

allows us, as usual, to state the general result of a measurement without assuming a particular outcome.[21]

The general result is that the logical entropy increase resulting from a measurement is the sum of the new distinction probabilities created by the join, which is the sum of the squared amplitudes of the off-diagonal indistinction amplitudes in the density matrix that were zeroed or "decohered" by the measurement.

In the example, the six off-diagonal amplitudes of $\frac{1}{3}$ were all zeroed so the change in logical entropy is: $6 \times \left(\frac{1}{3}\right)^2 = \frac{6}{9} = \frac{2}{3}$.

$$\rho(U) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \overset{measurement}{\longrightarrow} \hat{\rho}(U) = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.$$

[21] Note that this set-version of "decoherence" means actual reduction of state, not a "for all practical purposes" or FAPP [2] reduction.

In terms of sets, there are no distinctions in the indiscrete partition $\mathbf{0} = \{U\}$ so $h(\mathbf{0}) = \frac{|\mathrm{dit}(\mathbf{0})|}{|U \times U|} = 0$. Measurement by a non-degenerate attribute f gives the discrete partition $\mathbf{1} = f^{-1} \vee (U)$ where the distinctions are the ordered pairs (a, b), (a, c), and (b, c) together with the three opposite ordered pairs (b, a), (c, a), and (c, b) so the logical entropy is $h(\mathbf{1}) = \frac{|\mathrm{dit}(\mathbf{1})|}{|U \times U|} = \frac{6}{9} = \frac{2}{3}$. Those ordered pairs correspond exactly to off-diagonal terms zeroed in the transition $\rho(U) \to \hat{\rho}$ and $h(\mathbf{1}) = 1 - \mathrm{tr}\left[\hat{\rho}^2\right] = 1 - \left(\frac{1}{9} + \frac{1}{9} + \frac{1}{9}\right) = \frac{2}{3}$.

In this manner, the density matrices of QM/sets capture the set-based operations of logical information theory, and that, in turn, shows *how to interpret the density matrices of full QM* in terms of coherence and decoherence probabilities. The usual notion of von Neumann entropy in quantum information theory provides no such information-theoretic term-by-term interpretation of density matrices, not to mention of the process of measurement. In this manner, QM/sets shows, from the information-theoretic viewpoint, the essence at the logical level of what is going on in the full QM, i.e., QM/sets shows the "logic" of QM. The further development of the classical or quantum information theory using logical entropy is beyond the scope of this introductory paper [9].

11 Quantum computation theory in QM/sets

11.1 Qubits over 2 and non-singular gates

In QM over $\mathbb{C}$, a *quantum bit* or *qubit* is a non-zero (normalized) vector in $\mathbb{C}^2$. A standard orthonormal basis is denoted $|0\rangle$ and $|1\rangle$ so a qubit can be any (normalized) superposition $\alpha|0\rangle + \beta|1\rangle$ for $\alpha, \beta \in \mathbb{C}$. In QM/sets, i.e., QM over $\mathbb{Z}_2$, a *qubit over 2* or *qubit/2* is any non-zero vector in $\mathbb{Z}_2^2$ which for a given basis $|0\rangle$ and $|1\rangle$ would have the form $\alpha|0\rangle + \beta|1\rangle$ for $\alpha, \beta \in \mathbb{Z}_2$. As previously noted, Schumacher and Westmoreland (S&W) [21] restrict their treatment of Dirac's brackets to take values in the base field of $\mathbb{Z}_2$ which precludes a probability calculus so they develop a modal interpretation (0 = impossible and 1 = possible). Hence they call a non-zero vector in $\mathbb{Z}_2^2$ a "mobit" and call the resulting theory "modal quantum theory." Since our different treatment of the brackets yields a full probability calculus in QM/sets, we will not use the "modal" terminology but, nevertheless, their "mobit" is the same as our "qubit/2."

In $\mathbb{C}^2$, there is a continuum of qubits $\alpha|0\rangle + \beta|1\rangle$ for $\alpha, \beta \in \mathbb{C}$ but in $\mathbb{Z}_2^2$, there are only 3 qubits/2, namely $|0\rangle$, $|1\rangle$, and $|0\rangle + |1\rangle$. Hence a qubit/2 can be seen as the simplest possible extension beyond the classical bit with the two possibilities $|0\rangle$ and $|1\rangle$ by adding the superposition $|0\rangle + |1\rangle$.[22] As already noted in our treatment of Bell's Theorem in QM/sets, there are only three basis sets for $\mathbb{Z}_2^2$; any two of non-zero vectors are a basis with the third as their superposition.

In QM/sets (as in S&W's modal quantum theory), the dynamics are given by non-singular transformations which may be represented as non-singular zero-one matrices (which have non-zero determinants mod 2). A qubit over 2, $\alpha|0\rangle + \beta|1\rangle$, is represented in the standard basis $|0\rangle$ and $|1\rangle$ by the column vector $[\alpha, \beta]^t$.

The non-singular transformations are the *gates* that may be used in an algorithm for quantum computing over 2 (QC/2). The two one-qubit gates that carry over from quantum computing over $\mathbb{C}$ are the:

$$\textit{identity } I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \textit{negation } X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The four other one-qubit/2 gates in QC/2 are non-singular but when interpreted as matrices in $\mathbb{C}^2$ are not unitary. In particular, there is no requirement that a gate preserves the norm of a vector. One one-qubit/2 gate puts $|0\rangle$ into the superposition $|0\rangle + |1\rangle$ and leaves $|1\rangle$ the same:

[22] Here we are following the mild conceptual sloppiness common in the field of referring to any binary option as a "classical bit" when the bit as defined in Shannon's information theory is actually an *equiprobable* binary option. The comparable notion in logical information theory is a *distinction* or *dit* of a partition π on U which is exactly defined as an ordered pair (u, u') elements distinguished by π in the sense of the elements being in distinct blocks of π.

$$H_0 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Similarly another one-qubit/2 gate puts $|1\rangle$ into the superposition and leaves $|0\rangle$ the same:

$$H_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

And finally the other two one-qubit/2 gates are their negations:

$$XH_0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } XH_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

These six one-qubit/2 gates are the only non-singular transformations $\mathbb{Z}_2^2 \to \mathbb{Z}_2^2$.

As we will see below, some problems like the simplest Deutsch problem of determining if a single-variable Boolean function is balanced or constant can be solved in QC/2 solely with one-qubit/2 gates, whereas the usual solution to that problem in quantum computing over $\mathbb{C}$ uses two-qubit gates (four dimensional matrices). This is not as paradoxical as it may seem if we recall that quantum computing over 2 allows non-singular gates whereas the gates over $\mathbb{C}$ have to be unitary.

In representing these gates in the standard basis, we will use the standard Alice-Bob convention that the first or top one-qubit/2 (on the left) belongs to Alice and the second or bottom one-qubit/2 (on the right) belongs to Bob so the four basis vectors are: $|0_A\rangle\otimes|0_B\rangle = |0_A0_B\rangle$, $|0_A\rangle\otimes|1_B\rangle = |0_A1_B\rangle$, $|1_A\rangle \otimes |0_B\rangle = |1_A0_B\rangle$, and $|1_A\rangle \otimes |1_B\rangle = |1_A1_B\rangle$ (they are arranged in that order in the column vectors).

One two-qubit gate that carries over from quantum computing over $\mathbb{C}$ is the *controlled negation* gate:

$$Cnot_A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

which may be represented as acting on Alice's top line and Bob's bottom line:

$$\begin{array}{c} \to \bullet \to \\ \quad | \quad \\ \to \oplus \to \end{array}.$$

The action of a gate is specified by how it acts on the basis vectors. For either case $|0_A0_B\rangle$ and $|0_A1_B\rangle$ where Alice's qubit/2 is $|0_A\rangle$, the gate acts like the identity. But in the cases $|1_A0_B\rangle$ and $|1_A1_B\rangle$ where Alice's qubit/2 is $|1_A\rangle$, then Bob's qubit/2 is negated so that $|1_A0_B\rangle \to |1_A1_B\rangle$ and $|1_A1_B\rangle \to |1_A0_B\rangle$. In this case, Alice's qubit/2 is said to be the *controlling* qubit/2 (indicated by the subscript on $Cnot_A$) and Bob's the *target* qubit/2.

The controlling and target roles are reversed in the gate:

$$Cnot_B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \text{ represented as } \begin{array}{c} \to \oplus \to \\ \quad | \quad \\ \to \bullet \to \end{array}$$

where if Bob's qubit/2 is $|0_B\rangle$, then it acts like the identity, but if Bob's qubit/2 is $|1_B\rangle$, then Alice's qubit/2 is negated.

In a two-qubit/2 system, if a one-qubit/2 gate is to be applied to only one line, then tensor product of matrices is used. For instance to apply H_0 only to Bob's line, the two-qubit/2 gate is:

$$I \otimes H_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \text{ represented as } \rightarrow \boxed{\overrightarrow{H_0}} \rightarrow .$$

11.2 Teleportation of a qubit/2 with 1 classical bit

S&W's treatment [21] of the no-cloning theorem and superdense coding would work the same in QC/2 so we will not repeat it here. But after their treatment of superdense coding (of two classical bits), they remark: "The same set of entangled mobit states and single-mobit transformations can also be used to accomplish the MQT analogue of quantum teleportation." [21, p. 924] But that MQT (modal quantum theory) analogue of the usual quantum teleportation in full QM is somewhat odd since there are only three possible non-zero qubits/2 or mobits, and two classical bits suffice to transmit the identity of four different states–without entanglement having anything to do with it–if Alice knew which of the three mobits she had. It would be more in the spirit of quantum teleportation to transmit a qubit/2 (or mobit) using only one classical bit so that the entanglement has a real role. That is what we do.

In contrast with the usual two-bit teleportation scheme ([3], [20, pp. 26-28]), Alice only has one line instead of two, and she starts off with the qubit/2 $|\psi\rangle = \alpha|0_A\rangle + \beta|1_A\rangle$ to be teleported to Bob, while Bob starts with the usual $|0_B\rangle$, so the initial state in the two-qubit/2 system is $|\varphi_0\rangle = (\alpha|0_A\rangle + \beta|1_A\rangle) \otimes |0_B\rangle = \alpha|0_A0_B\rangle + \beta|1_A0_B\rangle$. The circuit diagram for the one-bit teleportation protocol is:

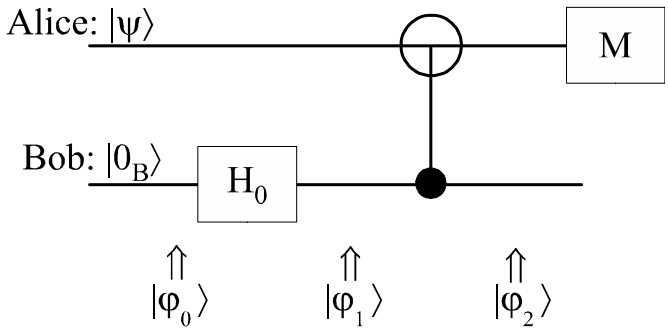

Figure 10: Teleportation scheme for a qubit/2 using 1 classical bit

where $\boxed{M}$ refers to Alice measuring her qubit/2. Alice and Bob start off together. First Bob applies the H_0 gate to his line (i.e., $I \otimes H_0$ is applied to both lines) to put Bob's state in the superposition $|0_B\rangle + |1_B\rangle$:

$$|\varphi_1\rangle = (I \otimes H_0) \begin{bmatrix} \alpha \\ 0 \\ \beta \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ 0 \\ \beta \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha \\ \beta \\ \beta \end{bmatrix}$$
$$= \alpha(|0_A0_B\rangle + |0_A1_B\rangle) + \beta(|1_A0_B\rangle + |1_A1_B\rangle)$$
$$= (\alpha|0_A\rangle + \beta|1_A\rangle) \otimes (|0_B\rangle + |1_B\rangle).$$

That non-entangled mutual state is then entangled by applying the $Cnot_B$ gate:

$$|\varphi_2\rangle = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \alpha \\ \beta \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ \beta \\ \alpha \end{bmatrix}$$

$$= \alpha\left(|0_A 0_B\rangle + |1_A 1_B\rangle\right) + \beta\left(|0_A 1_B\rangle + |1_A 0_B\rangle\right)$$
$$= |0_A\rangle \otimes (\alpha\,|0_B\rangle + \beta\,|1_B\rangle) + |1_A\rangle \otimes (\beta\,|0_B\rangle + \alpha\,|1_B\rangle).$$

Then Bob and Alice "separate" (like a pair of particles in the EPR experiment) so their only connection is the entangled state–and a classical communication channel for one classical bit. Without further operations, Alice then measures her line and gets either a $|0_A\rangle$ or $|1_A\rangle$. If she gets $|0_A\rangle$, then the state on Bob's line is $\alpha\,|0_B\rangle + \beta\,|1_B\rangle$ so that $|\psi\rangle = \alpha\,|0_A\rangle + \beta\,|1_A\rangle$ has been teleported to Bob. If Alice gets $|1_A\rangle$ then Bob's state is $\beta\,|0_B\rangle + \alpha\,|1_B\rangle$ so he only need apply the negation gate X to get $\alpha\,|0_B\rangle + \beta\,|1_B\rangle$. Hence Alice only has to send one classical bit with 0 = "do nothing" and 1 = "apply X" in order to tell Bob how to get the teleported state on his line. Taking M as the classical bit sent by Alice and $X^0 = I$, then the instruction to Bob is to apply X^M to his state to get the teleported state.

Replace the non-unitary but non-singular H_0 by the unitary *Hadamard matrix*

$$H = \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

and the protocol will teleport a full qubit $|\psi\rangle = \alpha\,|0_A\rangle + \beta\,|1_A\rangle \in \mathbb{C}^2$ with *one* classical bit. That (little known) protocol is called *X-teleportation*, was developed by Charles Bennett, and was analyzed, along with some other single-bit teleportation schemes, by Zhou, Leung, and Chuang [28].

11.3 Deutsch's simplest problem in QC/2

Deutsch's simplest problem is that of determining if a given Boolean function $y = f(x)$ is *balanced* in the sense of being one-one or is *constant* (two-to-one). An equivalent classification of the four unary Boolean functions is whether their *parity* in the sense of the mod 2 sum of their values $f(0) + f(1)$ is odd (balanced) or even (constant)–which is called the *parity satisfiability problem* or *Parity SAT* [25]. In the usual treatment of Deutsch's problem in quantum computation over $\mathbb{C}$, the gates U_f that evaluate the function are 4×4 gates which are unitary. But in quantum computing over 2, the gates need only be non-singular. A scheme to encode the four functions in non-singular evaluation 2×2 gates is:

$$E_f = X^{f(1)} H_{f(0)}$$

so the four function evaluation gates are:

$$f = X \text{ so } f(0) = 1 \text{ and } f(1) = 0\text{: } E_f = X^0 H_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} f(0) & f(1)+1 \\ f(1) & f(0) \end{bmatrix};$$
$$f = I \text{ so } f(0) = 0 \text{ and } f(1) = 1\text{: } E_f = X^1 H_0 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} f(1) & f(0)+1 \\ f(0)+1 & f(1)+1 \end{bmatrix};$$
$$f = 0 \text{ so } f(0) = 0 \text{ and } f(1) = 0\text{: } E_f = X^0 H_0 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} f(0)+1 & f(1) \\ f(1)+1 & f(0)+1 \end{bmatrix};$$
$$f = 1 \text{ so } f(0) = 1 \text{ and } f(1) = 1\text{: } E_f = X^1 H_1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} f(1)+1 & f(0) \\ f(0) & f(1) \end{bmatrix}.$$

Then it is evident that the mod 2 sum across the rows is the same for all four cases:

$$E_f \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} f(0) + f(1) + 1 \\ f(0) + f(1) \end{bmatrix}$$

so we only need measure that one-qubit/2 line to determine the function's parity. If the result is $|0\rangle$, then $f(0) + f(1) + 1 = 1$ (and $f(0) + f(1) = 0$) so the parity is even (or function is constant) and if the result is $|1\rangle$, then $f(0) + f(1) = 1$ so the parity is odd (or function is balanced). Hence the circuit diagram for the QC/2 algorithm is:

$$|0\rangle \longrightarrow \boxed{H_0} \longrightarrow \boxed{E_f} \longrightarrow \boxed{M}$$

QC/2 algorithm for the Deutsch problem or Parity SAT problem for unary Boolean functions

and the matrix operation giving the one-qubit/2 to be measured is:

$$X^{f(1)}H_{f(0)}H_0\begin{bmatrix}1\\0\end{bmatrix} = X^{f(1)}H_{f(0)}\begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = X^{f(1)}H_{f(0)}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}f(0)+f(1)+1\\f(0)+f(1)\end{bmatrix}.$$

This is the same Deutsch problem usually solved by a two-qubit circuit in full quantum computing over $\mathbb{C}$. In either case, two classical function evaluations are needed to determine the parity of the sum of the functions values so the "quantum speedup" is seen in the quantum algorithm in QC/2 or full QC only requiring one function evaluation.

11.4 The general Parity SAT problem solved in QC/2

The generalization to n-ary Boolean functions $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ is simple for the problem of determining the parity of the function where the parity is determined by the mod 2 sum of the function's 2^n values. To keep the notation manageable, we will consider the case of $n = 2$ which will make the pattern clear.

The function evaluation matrices E_f for binary Boolean functions $y = f(x_1, x_2)$ may be taken as:

$$E_f = X^{f(0,1)}H_{f(0,0)} \otimes X^{f(1,1)}H_{f(1,0)}.$$

Consider the binary Boolean function of the truth-functional conditional or implication $x_1 \Rightarrow x_2$ where (simplifying $f(0,0)$ to f_{00} etc.) $f_{00} = f_{01} = f_{11} = 1$ but $f_{10} = 0$, the function evaluation matrix is:

$$\begin{aligned}
&X^{f(0,1)}H_{f(0,0)} \otimes X^{f(1,1)}H_{f(1,0)} = X^1H_1 \otimes X^1H_0\\
&= \begin{bmatrix}0 & 1\\1 & 1\end{bmatrix} \otimes \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix} = \begin{bmatrix}f_{01}+1 & f_{00}\\f_{00} & f_{01}\end{bmatrix} \otimes \begin{bmatrix}f_{11} & f_{10}+1\\f_{10}+1 & f_{11}+1\end{bmatrix}\\
&= \begin{bmatrix}(f_{01}+1)\begin{bmatrix}f_{11} & f_{10}+1\\f_{10}+1 & f_{11}+1\end{bmatrix} & f_{00}\begin{bmatrix}f_{11} & f_{10}+1\\f_{10}+1 & f_{11}+1\end{bmatrix}\\ f_{00}\begin{bmatrix}f_{11} & f_{10}+1\\f_{10}+1 & f_{11}+1\end{bmatrix} & f_{01}\begin{bmatrix}f_{11} & f_{10}+1\\f_{10}+1 & f_{11}+1\end{bmatrix}\end{bmatrix}\\
&= \begin{bmatrix}(f_{01}+1)f_{11} & (f_{01}+1)(f_{10}+1) & f_{00}f_{11} & f_{00}(f_{10}+1)\\(f_{01}+1)(f_{10}+1) & (f_{01}+1)(f_{11}+1) & f_{00}(f_{10}+1) & f_{00}(f_{11}+1)\\f_{00}f_{11} & f_{00}(f_{10}+1) & f_{01}f_{11} & f_{01}(f_{10}+1)\\f_{00}(f_{10}+1) & f_{00}(f_{11}+1) & f_{01}(f_{10}+1) & f_{01}(f_{11}+1)\end{bmatrix}.
\end{aligned}$$

The key to any quantum algorithm is the clever use of superposition to extract the needed information. In this case, the superposition just adds up each row, which after some simplification, yields the two-qubit/2 column vector:

$$\begin{bmatrix}(f_{00}+f_{01}+1)(f_{10}+f_{11}+1)\\(f_{00}+f_{01}+1)(f_{10}+f_{11})\\(f_{00}+f_{01})(f_{10}+f_{11}+1)\\(f_{00}+f_{01})(f_{10}+f_{11})\end{bmatrix}.$$

And *regardless* of the binary function (calculated above for the implication), the above column vector of the row sums in terms of the function values is always the *same*![23] Note further that regardless of

[23] The proof is just an elaboration on the fact that the row sums of the tensor product of two matrices is the product of the row sums of the two matrices.

the values of the function, only one row has sum of 1 and the others sum to 0. Hence we only need to measure that two-qubit/2 to determine the parity of the function. Moreover the sum of values of the function occur in pairs with the first variable fixed, e.g., $f_{00}+f_{01}$ and $f_{10}+f_{11}$, so each row sum also contains the information about the parity of those unary functions $f(0,x_2)$ and $f(1,x_2)$. Hence the significance of the row sums being 1 is:

$$\begin{bmatrix} (f_{00}+f_{01}+1)(f_{10}+f_{11}+1) \\ (f_{00}+f_{01}+1)(f_{10}+f_{11}) \\ (f_{00}+f_{01})(f_{10}+f_{11}+1) \\ (f_{00}+f_{01})(f_{10}+f_{11}) \end{bmatrix} = \begin{bmatrix} 1=EE \\ 1=EO \\ 1=OE \\ 1=OO \end{bmatrix}.$$

For instance, if the measurement gives $|10\rangle$, then the entry in the third row $(f_{00}+f_{01})(f_{10}+f_{11}+1)$ is 1 which can only happen if each factor is 1 so the unary function $f(0,x_2)$ is odd and the unary function $f(1,x_2)$ is even.[24] Hence the 1 in the third row signifies $1=OE$. The parity of the whole binary function is immediately determined by the parity of those two unary functions since the sum of all the values is only even in the EE and OO cases (since the rule for adding even and odd numbers is: $E+E=E=O+O$), and is otherwise odd (since $E+O=O=O+E$).

Starting with the initial state $|00\rangle$, the gate $H_0\otimes H_0$ gives the superposition $[1,1,1,1]^t$ and the evaluation of the function gate E_f at that superposition takes the row sums of the evaluation matrix to yield the column vector to be measured. Hence the circuit diagram of two-qubit/2 gates is:

$$|00\rangle \longrightarrow \boxed{H_0\otimes H_0} \longrightarrow \boxed{E_f} \longrightarrow \boxed{M}$$

QC/2 algorithm for parity problem for binary Boolean functions.

This $n=2$ example indicates the pattern for the general case which uses n-qubit/2 gates in $\mathbb{Z}_2^2\otimes ...\otimes \mathbb{Z}_2^2$ (n times) $=\mathbb{Z}_2^{2^n}$:

$$|0...0\rangle \longrightarrow \boxed{H_0^{\otimes}n} \longrightarrow \boxed{E_f} \longrightarrow \boxed{M}$$

QC/2 algorithm for Parity SAT problem for n-ary Boolean functions.

The Unambiguous SAT problem is–when one is given or "promised" that a Boolean function has at most one case where it is satisfied–to find if it is satisfied or not. The solution to the Parity SAT problem also solves the Unambiguous SAT problem since "even" means no satisfying cases and "odd" means one satisfying case.[25]

The quantum speedup is particularly clear since classically each of the 2^n values of an n-ary Boolean function needs to be evaluated to determine the parity of the sum of the values, but the QC/2 algorithm only makes one functional evaluation for any n.

12 Concluding overview

QM/sets is the set version of the mathematics of quantum mechanics–without any specifically physical concepts (e.g., the Hamiltonian or DeBroglie relations). The connection between the two mathematical theories is the sets-to-vector-spaces bridge (or ladder) provided by the basis principle and used particularly by Weyl, but also by von Neumann and many others as it is essentially part of the mathematical folklore.

In the context of toy models of QM on vector spaces over finite fields (i.e., "modal quantum theory" [21], "discrete quantum theory" [14], "Galois field quantum mechanics" [5], or "mutant

[24] It could also be arranged for the pairs to represent the other two unary functions $f(x_1,0)$ and $f(x_1,1)$ by changing the functional evaluation matrix to: $X^{f(1,0)}H_{f(0,0)}\otimes X^{f(1,1)}H_{f(0,1)}$.

[25] Hanson et al. [14] give a somewhat more complicated algorithm that solves the Unambiguous SAT problem in QC/2.

quantum mechanics" [24]), the special case of the base field $\mathbb{Z}_2$ stands out since vectors can then be interpreted as a natural mathematical objects, i.e., sets. It is *only* this special case of base field $\mathbb{Z}_2$ that engages the sets-to-vector-spaces bridge of the lifting program. Thus the notion of a partition of a set lifts to a direct sum decomposition of a vector space, a numerical attribute on a set lifts to a linear operator on the space, the inverse-image set partition given by the numerical attribute lifts to the direct sum decomposition given by the eigenspaces of a (diagonalizable) linear operator, and so forth.

The set version of some QM concept, result, or model represents the simplest $\mathbb{Z}_2$-based essentials or "logic" of the matter, and in that old-fashioned sense, QM/sets is proposed as the "logic" of QM. Thus QM/sets is not only of pedagogical importance by showing the "distilled down" essential logic of the subject; it provides a treatment of many aspects of "quantum weirdness" using simple set concepts and thus it adds to the conceptual understanding (and demystification) of QM. For instance, the probability calculus of QM distills down in QM/sets to the usual Laplace-Boole calculus of logical finite probability theory (reformulated in a "non-commutative" fashion over the vector space $\mathbb{Z}_2^{|U|}$ which allows different bases instead of just one set U of outcomes). And quantum entanglement distills down in QM/sets to joint probability distributions on the direct product of two sets being correlated rather than independent, and Bell's Theorem carries over to sets by showing that the probabilities involved in QM/sets measurements could not come from an independent joint distribution.

Quantum information theory based on QM/sets is essentially the logical information theory defined by the normalized counting measure on partitions (represented as partition relations or apartness relations) just as logical probability theory is defined by the normalized counting measure on subsets (events) of a universe set of outcomes. The normalized counting measure on partitions is the notion of logical entropy [9] that provides a new logical foundation for information theory. Shannon's notion of entropy is a higher-level concept adapted to the theory of communication (as Shannon always named the theory [22]). The notion of logical entropy of a partition can be formulated in terms of "delifted" density matrices and it provides an exact interpretation of the entries in a density matrix in terms of indistinction probabilities (and in terms of "coherence probabilities" in the relifted version). The Shannon notion of entropy lifted to quantum information theory as von Neumann entropy provides no such logical analysis of a state (pure or mixed) represented in a density matrix; it is suited for analyzing the quantum communications protocols lifted to QM from Shannon's theory of communication through classical channels.

In quantum computing in QM/sets or QC/2, the coefficients in the gates are only from $\mathbb{Z}_2$ but the gates need only be non-singular (unitarity is only defined on inner product spaces and vector spaces over finite fields have no inner products). In addition to the no-cloning theorem and superdense coding, there is a simple protocol for teleporting a qubit/2 using only one classical bit (that foreshadows a little-known single-bit protocol in full QM [28]). As an example of a quantum computing algorithm over 2 or $\mathbb{Z}_2$, the simplest Deutsch problem is reformulated as the Parity SAT problem for unary Boolean functions, and is solved by a simple algorithm using only one-qubit/2 gates. This then generalizes immediately to a QC/2 algorithm solving the general Parity SAT problem (to determine the parity of the sum of values of an n-ary Boolean function). As expected, the algorithm is so simple that the key role of superposition is obvious, and that superposition gives the quantum speedup of a single function evaluation in contrast with the 2^n evaluations needed classically. This shows that the quantum speedup has nothing to do with the greater power of calculations in the complex numbers $\mathbb{C}$ as opposed to classical computing using $\mathbb{Z}_2$, since QC/2 is also restricted to $\mathbb{Z}_2$.

Finally, quantum mechanics over sets or QM/sets is part of a research program that arose out of the recent development of the logic of partitions ([10] and [12]), the logic that is mathematically dual to the ordinary Boolean logic of subsets (usually mis-specified as the special case of "propositional" logic). This research program ultimately aims to interpret quantum mechanics using the notion of objective indefiniteness [11]. Quantum mechanics over sets is a key part of that program since

the fundamental QM notions such as: (1) eigenstates, (2) superpositions of eigenstates, and (3) measurement in vector spaces, are distilled down respectively into the set concepts: (1′) the "definite" singleton subsets $\{u\} \subseteq S$, (2′) the "indefinite" multiple-element subsets S that "superpose" (i.e., collect together) a number of definite elements, and (3′) the partition-join-action of a set partition, the inverse-image $\left\{f^{-1}(r)\right\}_r$ of a numerical attribute $f : U \to \mathbb{R}$, on a "pure" indefinite subset S to create a probabilistic "mixed state" $\left\{f^{-1}(r) \cap S\right\}_r$ of more definite subsets with probabilities $\Pr(r|S) = \frac{\left|f^{-1}(r) \cap S\right|}{|S|}$.

References

[1] Bell, John S. 1987. *Speakable and unspeakable in quantum mechanics*. Cambridge UK: Cambridge University Press.

[2] Bell, John S. 1990. Against "Measurement". In *Sixty-Two Years of Uncertainty*. Arthur I. Miller ed., New York: Plenum Press: 17-31.

[3] Bennett, C.H., G. Brassard, C. Crepeau, R. Jozsa, A. Peres and W. Wootters 1993. Teleporting an unknown quantum state via dual classical and EPR channels. *Phys. Rev. Lett.,* 70: 1895-1899.

[4] 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.

[5] Chang, Lay Nam, Zachary Lewis, Djordje Minic and Tatsu Takeuchi 2013. Galois Field Quantum Mechanics. *[quant-ph] arXiv:1205.4800v2.*

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

[7] D'Espagnat, Bernard 1979. The quantum theory and reality. *Scientific American*. 241 (5): 158-181.

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

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

[10] 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. Downloadable at: www.ellerman.org.

[11] Ellerman, David 2013. The Objective Indefiniteness Interpretation of Quantum Mechanics, *[quant-ph] arXiv:1210.7659.*

[12] Ellerman, David (forthcoming). An Introduction to Partition Logic. *Logic Journal of the IGPL.*

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

[14] Hanson, Andrew J., Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai 2013. *Discrete Quantum Theories*. arXiv:1305.3292v1: 13 pages.

[15] Hardy, L. 1993. Non-locality for two particles without inequalities for almost all entangled states. *Phys. Rev. Lett.*, 71, 1665.

[16] Hoffman, Kenneth and Ray Kunze 1961. *Linear Algebra*. Englewood Cliffs NJ: Prentice-Hall.

[17] Hughes, R.I.G. 1989. *The structure and interpretation of quantum mechanics*. Cambridge: Harvard University Press.

[18] Jammer, Max 1974. *The Philosophy of Quantum Mechanics: The Interpretations of Quantum Mechanics in Historical Perspective*. New York: John Wiley.

[19] 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.

[20] Nielsen, Michael and Isaac Chuang 2000. *Quantum Computation and Quantum Information*. Cambridge UK: Cambridge University Press.

[21] Schumacher, B. and M. Westmoreland 2012. Modal Quantum Theory. *Foundations of Physics*, 42, 918-925.

[22] Shannon, Claude E. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal*. 27: 379-423; 623-56.

[23] Shimony, Abner 1986. Events and processes in the quantum world. In *Quantum Concepts in Space and Time*. R. Penrose and C. Isham ed., Oxford: Oxford University Press: 182-203.

[24] Takeuchi, Tatsu, Lay Nam Chang, Zachary Lewis and Djordje Minic 2012. Some Mutant Forms of Quantum Mechanics. *[quant-ph] arXiv:1208.5544v1*.

[25] Valiant, L. G. and V. V. Vazirani 1986. NP is as easy as detecting unique solutions. *Theoretical Computer Science*. 47: 85-93.

[26] von Neumann, John 1955. *Mathematical Foundations of Quantum Mechanics*. Robert T. Beyer trans., Princeton NJ: Princeton University Press.

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

[28] Zhou, Xinlan, Debbie W. Leung and Isaac L. Chuang 2000. Methodology for quantum logic gate construction. *[quant-ph] arXiv:0002039v2*. 17 pages.